

**QUERY EXPANSION PADA SISTEM TEMU KEMBALI  
INFORMASI DENGAN MODEL RUANG VEKTOR**

**TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat  
Untuk Memperoleh Gelar Sarjana Teknik Pada  
Jurusan Teknik Informatika

Oleh :

**KHAIRI LESTARI**

**10851001784**



**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU  
PEKANBARU**

**2013**

# **LEMBAR PERSETUJUAN**

## ***QUERY EXPANSION* PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MODEL RUANG VEKTOR**

### **TUGAS AKHIR**

oleh:

**KHAIRI LESTARI**  
**10851001784**

Telah diperiksa dan disetujui sebagai laporan tugas akhir  
di Pekanbaru, pada tanggal 11 Februari 2013

Koordinator Tugas Akhir

Pembimbing

**Iwan Iskandar, M.T**  
**NIK. 130 508 071**

**Surya Agustian, S.T, M.Kom**  
**NIP. 19760830 201101 1 003**

## LEMBAR PENGESAHAN

### ***QUERY EXPANSION* PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MODEL RUANG VEKTOR**

#### TUGAS AKHIR

oleh :

**KHAIRI LESTARI**  
**10851001784**

Telah dipertahankan di depan sidang dewan penguji  
Sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
Di Pekanbaru, pada tanggal 11 Februari 2013

Pekanbaru, 11 Februari 2013

Mengesahkan

Dekan

Ketua Jurusan

**Dra. Hj. Yenita Morena, M.Si**  
**NIP. 19601125 198503 2 002**

**Novriyanto S.T, M.Sc**  
**NIP. 19771128 200710 1 003**

#### **DEWAN PENGUJI**

Ketua : Surya Agustian, S.T, M.Kom \_\_\_\_\_

Sekretaris : Surya Agustian, S.T, M.Kom \_\_\_\_\_

Anggota I : Iwan Iskandar, M.T \_\_\_\_\_

Anggota II : M. Safrizal, S.T, M.Cs \_\_\_\_\_

## **LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL**

Tugas akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh tugas akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan tugas akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Pekanbaru, 11 Februari 2013

**KHAIRI LESTARI**

## **LEMBAR PERNYATAAN**

Dengan ini saya menyatakan bahwa dalam tugas akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 11 Februari 2013

Yang membuat pernyataan,

**KHAIRI LESTARI**

## LEMBAR PERSEMBAHAN



“ Tiada daya upaya dan kekuatan, kecuali dengan pertolongan Allah Yang Maha Tinggi lagi Maha Agung. Maha Suci Allah, yang ditangan-Nyalah segala kerajaan dan Dia Maha Kuasa atas segala sesuatu”.  
(QS.Al-Mulk (67))

**Kebahagiaan, kedamaian, dan ketenraman hati senantiasa berawal dari ilmu pengetahuan. Itu karena ilmu mampu menembus yang samar, menemukan sesuatu yang hilang, dan menyingkap yang tersembunyi. serta, naluri dari jiwa.**

**Kebodohan itu sangat membosankan dan menyedihkan. karena ia tidak pernah memunculkan hal baru yang lebih menarik dan segar yang kemarin seperti hari ini, dan yang hari ini pun akan sama dengan yang akan terjadi esok hari.**

**Bila anda ingin senantiasa bahagia, tuntutlah ilmu, galilah pengetahuan, dan raihlah berbagai manfaat, niscaya semua kesedihan, kepedihan, dan kecemasan itu akan sirna.**

**Janganlah seseorang sombong dengan harta dan kedudukannya, kalau memang ia tak memiliki ilmu sedikit pun. Sebab, kehidupannya tidak akan sempurna.**

**Dr. Aidh Al – Qarni**

Assalamu'alaikum Wr.Wb

Dan rasa syukur yang tak terkira dari hamba-Mu Ya Allah dan baginda rasul Muhammad Saw

Alhamdulillahirobbil'alamiin..

Do'a-mu dan ridho-mulah yang selalu menyertaiku dan memberikan jalan yang lebih mudah dalam setiap langkahku. Semua ini ananda persembahkan untuk Ayahanda, Ibunda tercinta terima kasih atas segala kasih sayang, doa, didikan, dan pengorbanan yang telah engkau lakukan untuk ananda, untuk kaka tersayang Rahmi lestari yang tidak bosan-bosannya memberikan nasehat dukungan dan semangat. Untuk adik-adiku tersayang Rifki, Kiki dan Sena. Untuk Keluarga Besarku, Teman Seperjuangan TIF o8 A, dan buat semua teman-temanku tanpa terkecuali Kalian lah semangat hidupku dan sumber kebahagiaanku, Terima Kasih atas kebersamaan yang kita lalui selama ini, baik dalam duka maupun suka..

**11 Februari 2013**

# ***QUERY EXPANSION* PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MODEL RUANG VEKTOR**

**KHAIRI LESTARI**

**10851001784**

Jurusan Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau

## **ABSTRAK**

Tidak maksimalnya STBI (Sistem temu kembali informasi) jika *query* yang digunakan sangat pendek hal ini dikarenakan kerja dari STBI hanya pada pada tingkat kata atau kalimat saja dan tidak sampai pada pergeseran makna pada *query* tersebut, oleh karena itu untuk mengatasi permasalahan ini diusulkanlah perluasan pada *query* yang ada dengan menambahkan kata perluasan dari *query*, penelitian ini bertujuan untuk membandingkan hasil pencarian sebelum dilakukan perluasan dan setelah dilakukan perluasan pada *query* pencarian, STBI yang dibangun menggunakan model ruang vektor dan perluasan *query* menggunakan kamus tesaurus bahasa Indonesia. Dan dengan menggunakan pengujian *recall* dan *precision* terhadap performa hasil pencarian maka didapatkan hasil penelitian ini, ternyata hasil pencarian sistem temu kembali informasi dengan model ruang vektor tanpa perluas tingkat relevansinya lebih bagus jika dibandingkan dengan hasil pencarian setelah diperluas dan dengan perbandingan persentasi sebesar 57% jika diperluas sedangkan setelah diperluas menjadi 47% hal ini dikarenakan hasil pencarian setelah *query* diperluas sistem hanya mampu mengembalikan dokumen yang tidak relevan terhadap *query*.

**Kata kunci:** Model Ruang Vektor, *Query Expansion*, *Recall* dan *Precision*, Sistem Temu Kembali Informasi, Tesaurus.

# ***QUERY EXPANSION ON INFORMATION RETRIEVAL SYSTEM WITH VECTOR SPACE MODEL***

**KHAIRI LESTARI**

**108510017784**

*Information Engineering Department  
Faculty of Sciences and Technology  
State Islamic University of Sultan Syarif Kasim Riau*

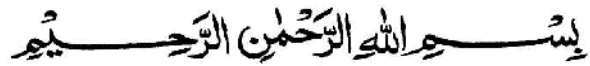
## ***ABSTRACT***

*Not maximum IRS (information retrieval system) if the query be used is very short, because work of IRS only at level of word or sentence and not come to meaning of the query proficiency level, therefore to overcome this problem proposed expansion on query by adding from query, this research aimed to compare of result before using query expansion and after using query expansion, IRS built using the vector space model and query expansion using word from thesaurus Indonesian. With use the test recall and precision performance of the result, then result if this research where information retrieval system with vector space model without expansion the relevant level more then compared with the result after expansion where percentage ratio 57% while after expansion percentage ratio only 47%. This is because the search result after query expansion only able to return documents that are not relevant to the query.*

***Key words :****Information Retrieval Systems, Vector Space Model, Query Expansion, Recall and Precision, Thesaurus*



## KATA PENGANTAR



*Assalammu'alaikum wa rahmatullahi wa barakatuh.*

*Alhamdulillah rabbil'alamin*, , puji dan syukur penulis ucapkan kepada Allah SWT atas segala rahmat dan hidayah yang diberikan-Nya, sehingga penulis dapat melaksanakan dan menyelesaikan tugas akhir ini dengan baik. Shalawat beriring salam diucapkan untuk junjungan kita Rasulullah Muhammad SAW, karena jasa Beliau kita bisa menikmati zaman yang penuh dengan ilmu pengetahuan seperti sekarang ini.

Tugas akhir yang berjudul **QUERY EXPANSION PADA SISTEM TEMU KEMBALI INFORMASI DENGAN MODEL RUANG VEKTOR** ini disusun sebagai satu syarat untuk mendapatkan gelar kesarjanaan pada jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.

Dalam penulisan dan penyusunan laporan tugas akhir ini penulis tidak terlepas dari bantuan berbagai pihak yang telah memberikan masukan berupa kritik, saran, motivasi dan dorongan yang sangat bermanfaat bagi penulis. Untuk itu dalam kesempatan ini penulis menyampaikan terima kasih yang tak terhingga kepada :

1. Allah SWT, yang dengan rahmatNya memberikan semua yang terbaik dan yang dengan hidayahNya memberikan petunjuk sehingga dalam penyusunan laporan ini berjalan lancar.
2. Rasulullah SAW, yang telah membawa petunjuk bagi manusia agar menjadi manusia paling mulia derajatnya di sisi Allah SWT.
3. Bapak Prof. DR. H.M. Nazir, selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Dra. Hj. Yenita Morena, M.Si, selaku Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau.

5. Bapak Novriyanto, S.T, M.Sc, selaku Ketua Jurusan Teknik Informatika, Fakultas Sains dan Teknologi.
6. Bapak Febiyanto, S.T, M.Kom selaku Pembimbing Akademis Penulis.
7. Bapak Iwan Iskandar, M.T sebagai koordinator tugas akhir yang telah memberi masukan-masukan untuk penyelesaian tugas akhir ini.
8. Bapak Surya Agustian, S.T, M.Kom selaku Pembimbing tugas akhir yang selalu sabar dan meluangkan waktu untuk membimbing, memberikan saran, kritik, ilmu, dukungan, dan motivasinya yang luar biasa dalam penyusunan tugas akhir ini.
9. Bapak Iwan Iskandar, M.T selaku penguji I, dan Bapak M.Safrizal, S.T,M.Cs Selaku penguji II, yang telah banyak memberikan masukan kepada penulis agar Tugas Akhir ini dapat selesai dengan baik dan lancar.
10. Bapak Benny Sukma Negara, M.T, Ibu Rizqa Raaiqa B, ST dan Ibu Sonya Meitracie, ST yang telah sudi meluangkan waktunya untuk menjadi dosen Ahli didalam mengoreksi hasil pencarian dari sistem yang penulis buat.
11. Seluruh dosen dan staf Fakultas Sains dan Teknologi khususnya pada Jurusan Teknik Informatika. Terima kasih atas ilmu yang telah diberikan.
12. Kedua Orang Tua tercinta yakni bapak M.LAMAZI S.Pdi & Ibunda SABARIYAH, yang telah mendo'akan dan memberikan dukungan yang sangat luar biasa kepada penulis dalam menyelesaikan tugas akhir ini. Semoga beliau selalu dalam lindungan Allah SWT serta segala ketulusan dan pengorbanan beliau di ridhoi oleh Allah SWT.
13. Kaka Tercinta Rahmi Lestari S.Pd dan Abang Mukhlis yang selalu memberikan dukungan yang sangat besar terhadap penulis, yang selalu bisa mengerti keadaan penulis dan selalu menjadi panutan bagi penulis, serta adik-adiku yang telah menjadi motivasu Hidup untuk menjadi seorang kaka yang Baik, Riki Kiki dan Sena.
14. Seluruh keluarga besar Penulis yang berada di Sei Bela dan di Tembilahan. Nenek, Julak, Aci Imang, Aci Wati, Aci Iyang, Aci Yen, Aci

Suai, Aci Ijol, Aci Icai, Aci Icit dan busu Rido. Terimakasih telah menjadi keluarga yang selalu mengerti akan Hidup penulis.

15. Pihak-pihak yang berperan penting didalam membantu penulis didalam menyelesaikan penelitian ini yang telah sudi meluang kan waktunya dan membantu tanpa pamrih R.Syahroni, Fritayola, Novreni, Fauzi Azis, Arita, Rido M, Indah I dan Rendra dinata.
16. Teman-temanku seperjuangan Rendra, Rusdi, Fauzi, Fadli, Budi, Didi, Inop, Ali, Bg Yudi Emka, Asep, Ade, Darni, Novreni, Ulfi, Zulfa, Indah, Imel, Desi, Dewi, Robi Lisfi, Ridho, Gusman, Roni, Ilyas, Surya, Alimin, Robi Hendri, Abdi, Endriko, Eko Kesuma, Dani, Verdy, Yola, Yusuf Amirat, kristiawan, Rido, Adek R, Lia, Ika dan seluruh teman-teman angkatan TIF 08. Semoga kita bisa mencapai cita – cita kita semua.
17. Penghuni Kos Lanang SEJATI Edi, Aji, Jamil, Bohar, Mas Iwan dan Reno yang selalu siap untuk memberikan semangat bagi penulis disaat dibutuhkan.
18. Seluruh pihak yang belum penulis cantumkan, terima kasih atas dukungannya, baik material maupun spiritual.

Penulis menyadari bahwa dalam penulisan laporan ini masih banyak kesalahan dan kekurangan, oleh karena itu kritik dan saran yang sifatnya membangun sangat penulis harapkan untuk kesempurnaan laporan ini. Akhirnya penulis berharap semoga laporan ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya. Amin

*Wassalamu'alaikum wa rahmatullahi wa barakatuh*

Pekanbaru, 11 Februari 2013

**KHAIRI LESTARI**

## DAFTAR ISI

HALAMAN JUDUL LAPORAN .....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN .....	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN .....	v
LEMBAR PERSEMBAHAN .....	vi
ABSTRAK .....	vii
<i>ABSTRACT</i> .....	viii
KATA PENGANTAR .....	ix
DAFTAR ISI .....	xii
DAFTAR GAMBAR .....	xvi
DAFTAR TABEL.....	xviii
DAFTAR RUMUS .....	xx
DAFTAR ISTILAH .....	xxi
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	I-1
1.1. Latar Belakang .....	I-1
1.2. Rumusan Masalah .....	I-2
1.3. Batasan Masalah.....	I-2
1.4. Tujuan Penelitian .....	I-3
1.5. Sistematika Penulisan .....	I-3
BAB II LANDASAN TEORI.....	II-1
2. 1. Sistem Temu Kembali Informasi .....	II-1
2.1.1. Pengertian Sistem Temu Kembali Informasi .....	II-1
2.1.2. Tujuan dan Fungsi.....	II-2

2.1.3. Arsitektur Sistem Temu Balik Informasi .....	II-3
2.1.3.1. Koleksi Dokumen .....	II-4
2.1.3.2. <i>Query</i> .....	II-4
2.1.3.3. Inverter <i>Index</i> Berbobot .....	II-5
2.1.3.4. Membandingkan <i>Query</i> dengan Dokumen .....	II-6
2.1.4 Model Dalam Sistem Temu Bali Informasi .....	II-6
2.1.4.1. Model <i>Boolean</i> .....	II-6
2.1.4.2. Model Probabilistik .....	II-7
2.1.4.3 Model Ruang <i>Vektor</i> .....	II-8
2.1.5. Pengujian Sistem Temu Kembali Informasi .....	II-19
2.2. Tesaurus .....	II-19
2.3. <i>Query Expansion</i> .....	II-20
BAB III METODOLOGI PENELITIAN .....	III-1
3.1. Pemilihan Lokasi Penelitian .....	III-2
3.2. Pengumpulan Data .....	III-2
3.3. Identifikasi dan perumusan Masalah .....	III-2
3.4. Analisa Sistem .....	III-3
3.5. Perancangan Sistem .....	III-4
3.6. Implementasi .....	III-4
3.7. Pengujian Sistem .....	III-5
3.8. Kesimpulan dan Saran .....	III-5
BAB IV ANALISA DAN PERANCANGAN .....	IV-1
4.1. Analisa Sistem Lama .....	IV-1
4.2. Analisa Pengembangan Sistem .....	IV-1
4.2.1. Analisa Subsistem Dialog .....	IV-3
4.2.2. Analisa Subsistem Data .....	IV-6
4.2.3. Analisa Model .....	IV-7
4.2.3.1. Pembentukan Korpus .....	IV-7
4.2.3.2. Proses <i>Query</i> .....	IV-12

4.2.3.3. Perluasan <i>Query</i> .....	IV-12
4.2.3.4. Pencocokan ( <i>Similarity</i> ).....	IV-14
4.3. Analisa Perancangan Sistem .....	IV-16
4.3.1 Perancangan Database.....	IV-16
4.3.2. Perancangan Tampilan .....	IV-17
4.3.2.1. Form Tampilan Utama.....	IV-18
4.3.2.2. Form <i>Login</i> Admin .....	IV-18
4.3.2.3. Form Halaman Utama admin.....	IV-19
BAB V IMPLEMENTASI DAN PENGUJIAN .....	V-1
5.1. Implementasi .....	V-1
5.1.1. Batasan Implementasi .....	V-1
5.1.2. Lingkungan Operasional .....	V-2
5.1.3. Hasil Implementasi Sistem.....	V-2
5.1.3.1 Implementasi Perhitungan Sistem .....	V-2
5.1.3.2 Implementasi <i>Interface</i> Sistem .....	V-6
5.2. Pengujian.....	V-13
5.2.1. Ruang Lingkup Pengujian.....	V-14
5.2.2. Rencana Pengujian .....	V-14
5.2.3. Hasil Pengujian .....	V-14
5.2.3.1 Pengujian <i>Query</i> 1 .....	V-15
5.2.3.2 Pengujian <i>Query</i> 1 dengan Perluasan.....	V-17
5.2.3.3 Pengujian <i>Query</i> 2.....	V-19
5.2.3.4 Pengujian <i>Query</i> 2 dengan Perluasan.....	V-20
5.2.3.5 Pengujian <i>Query</i> 3.....	V-23
5.2.3.6 Pengujian <i>Query</i> 3 dengan Perluasan.....	V-25
5.2.4. Kesimpulan Pengujian.....	V-29
BAB VI PENUTUP .....	VI-1
6.1. Kesimpulan.....	VI-1

6.2. Saran.....	VI-2
DAFTAR PUSTAKA .....	xxv
LAMPIRAN	
RIWAYAT HIDUP	

## DAFTAR GAMBAR

Gambar	Halaman
2.1. Bagian-bagian Sistem Temu Balik Informasi .....	II-2
2.2. Arsitektur Sistem Temu Balik Informasi .....	II-3
2.3. Besar sudut antara vektor <i>query</i> dan vektor dokumen .....	II-9
2.4. Representasi Dokumen dan <i>Query</i> pada Model Ruang Vektor .....	II-10
2.5. Representasi Grafis Sudut Vektor Dokumen dan <i>Query</i> .....	II-13
3.1. <i>Flowchart</i> Penyusunan Tugas Akhir .....	III-1
4.1. <i>Flowchart</i> Sistem Secara Umum .....	IV-2
4.2 Konteks Diagram STBI QE .....	IV-3
4.3. DFD Level 1 .....	IV-4
4.4. DFD Level 2 Proses 1 .....	IV-5
4.5. ERD .....	IV-6
4.6. Tahap Proses <i>Query</i> Pengguna .....	IV-12
4.7. Tahap Proses <i>Query</i> Pengguna Dengan menerapkan Perluasan .....	IV-13
4.8. Analisa Form Tampilan Utama .....	IV-18
4.9. Analisa Form <i>Login</i> Admin .....	IV-18
4.10. Analisa Form Halaman Utama Admin .....	IV-19
4.11. Analisa Form Halaman Utama Admin .....	IV-20
5.1. Menu Utama Admin .....	V-6
5.2. <i>Form</i> Koleksi Dokumen Sistem .....	V-7
5.3. Hasil Input Koleksi Dokumen .....	V-7
5.4. Menampilkan Koleksi Dokumen Sistem .....	V-8
5.5. Menampilkan Hasil Pengindekan .....	V-8
5.6. <i>Form</i> Daftar <i>Stopword</i> .....	V-9
5.7. Menampilkan Lis karakter Yang Telah Diinputkan .....	V-9
5.8. Form Daftar <i>Stopword</i> .....	V-10
5.9. Menampilkan Lis Karakter Yang telah diinput .....	V-10



5.10. <i>Form</i> Daftar <i>Stop Word</i> .....	V-11
5.11. Menampilkan Lis pengguna yang telah diinputkan .....	V-11
5.12. <i>Form</i> Pencarian STBI.....	V-11
5.13. Hasil Pencarian .....	V-12
5.14. Hasil Pencarian Selengkapnya.....	V-12
5.15. Pengujian Hasil <i>Query</i> <sub>1</sub> .....	V-16
5.16. Pengujian Pengujian Q <sub>1</sub> Dengan Perluasan.....	V-17
5.17. Interpolasi <i>Recall</i> dan <i>Precision</i> Q <sub>1</sub> .....	V-19
5.18. Hasil Pengujian Q <sub>2</sub> .....	V-20
5.19 Hasil Pengujian Q <sub>2</sub> Dengan Perluasan .....	V-22
5.20. Interpolasi <i>Recall</i> dan <i>Precision</i> Q <sub>2</sub> .....	V-23
5.21 Hasil Pengujian Q <sub>3</sub> .....	V-24
5.22 Hasil Pengujian Q <sub>3</sub> Dengan Perluasan .....	V-26
5.20 Interpolasi <i>Recall</i> dan <i>Precision</i> Q <sub>3</sub> .....	V-27

## DAFTAR TABEL

Tabel	Halaman
2.1 Hasil Pembobotan Dokumen Model Ruang Vektor .....	II-16
2.2 Hasil Pembobotan <i>Query</i> Model Ruang Vektor .....	II-17
2.3 Contoh Asumsi dari <i>Precision</i> dan recall.....	II-19
2.4 Pengindekan Dokumen Serta perhitungan Bobot Dokumen.....	II-21
2.5 Perhitungan Bobot <i>Query</i> .....	II-23
2.6. Perhitungan Simmilarity (dj,q) Vektor .....	II-23
4.1. Keterangan Context Diagram .....	IV-3
4.2. Keterangan Dfd Level 1 .....	IV-4
4.3. Keterangan Dfd Level 2 Proses 1 .....	IV-5
4.4. ERD Sistem .....	IV-6
4.5. Penyimpanan kata kedalam <i>database</i> .....	IV-10
4.6. Perhitungan Bobot Kata .....	IV-11
4.7. Pembobotan <i>Query</i> .....	IV-12
4.8. Ilustrasi Perluasan <i>Query</i> .....	IV-13
4.9. Ilustrasi Pembobotan <i>query</i> yang diberi faktor penyesuaian.....	IV-14
4.10. Nilai Similarity Setiap dokumen terhadap <i>query</i> .....	IV-15
4.11. Bobot dari <i>query</i> hasil perluasan.....	IV-15
4.12. Nilai Similarity Setiap dokumen terhadap <i>query</i> setelah diperluas.....	IV-16
4.13. Pengguna .....	IV-16
4.14. <i>Index</i> .....	IV-16
4.15. Karakter .....	IV-17
4.16. <i>Stopword</i> .....	IV-17
4.17. Vektor dokumen.....	IV-17
4.18. Tesaurus.....	IV-17
4.19. Keterangan Struktur Menu Admin.....	IV-20
5.1. Hasil pencarian setelah diperluas.....	V-6

5.2. Daftar <i>query</i> Pengujian.....	V-15
5.3. Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_1$ ) Tanpa Perluasan .....	V-16
5.4. .Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_1$ ) Dengan Perluasan.....	V-18
5.5. Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_2$ ) Tanpa Perluasan.....	V-20
5.6. Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_2$ ) Dengan Perluasan .....	V-21
5.7. Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_3$ ) Tanpa Perluasan.....	V-23
5.8 Perhitungan <i>recall</i> dan <i>Presicion</i> ( $Q_3$ ) Dengan Perluasan .....	V-25
5.9. Hasil Pengujian $Q_1$ Tanpa Perluasan.....	V-26
5.10. Hasil Pengujian $Q_1$ Dengan Perluasan.....	V-28
5.11 Hasil Pengujian pada $Q_2$ Tanpa Perluasan.....	V-27
5.12 Hasil Pengujian pada $Q_2$ Dengan Perluasan .....	V-28
5.13. Hasil Pengujian pada $Q_3$ Tanpa Perluasan.....	V-28
5.14. Hasil Pengujian) pada $Q_3$ Dengan Perluasan .....	V-29

## DAFTAR RUMUS

Rumus	Halaman
2.1. <i>LOGARITHMIC TF</i> .....	II-11
2.2. <i>AUGMENTED TF</i> .....	II-12
2.3. <i>IDF</i> .....	II-12
2.4. <i>BOBOT</i> .....	II-13
2.5. <i>INNER PRODUCT</i> .....	II-14
2.6. <i>RELEVANSI</i> .....	II-14
2.7. <i>SIMILARITY</i> .....	II-15
2.8. <i>PRECISION</i> .....	II-19
2.9. <i>RECALL</i> .....	II-19
2.8. <i>PEMBOBOTAN QUERY EXPANSION</i> .....	II-21

## DAFTAR ISTILAH

<i>Vector Space Model</i>	=	model ruang vektor
<i>Query</i>	=	kata yang digunakan oleh pengguna untuk mewakili informasi yang ingin didapatkan.
<i>Query expansion</i>	=	Memperluas atau memperbannya kata awal berdasarkan sinonim dari kata yang ada didalam kamus.
<i>Term</i>	=	kata atau <i>query</i>
<i>Information retrieval</i>	=	Temu kembali informasi
<i>Sinonim</i>	=	Lawan Kata
<i>Retriev</i>	=	menemukan kembali
<i>Text Operation</i>	=	Operasi atau Preoses terhadap Teks
<i>Term Selection</i>	=	Pemilihan Kata
<i>Term Index</i>	=	Pemilihan Kata yang mewakili Dokumen
<i>Termed word</i>	=	Pemotongan Kata Tunggal
<i>Lower case</i>	=	Huruf kecil
<i>Filtration</i>	=	Penyaringan
<i>Stop-word</i>	=	Kata yang dihapus karna dianggap sering muncul dalam sebuah dokumen
<i>Stop-list</i>	=	Daftar dari kata <i>Stop-word</i>
<i>Retrieval function</i>	=	Fungsi Pencarian
<i>Relevance Notation</i>	=	Notasi Kesesuayan
<i>Boolean Model</i>	=	Salah Satu Model yang ada didalam <i>Information Retrieval</i> yang disebut dengan Model Bolean
<i>Probabilistic Model</i>	=	Salah Satu Model yang ada didalam <i>Information Retrieval</i> yang disebut dengan Model Bolen
<i>Proximity Operator</i>	=	Salahsatu cara Perlasan dari model Boolean
<i>Wildcard Operator</i>	=	Salahsatu cara Perlasan dari model Boolean
<i>Software</i>	=	Perangkat Lunak yang bisa dipasangkan kedalam Perangkat Elektronik
<i>Partial Matching</i>	=	Sebagian Sesuai

<b><i>Probabilistic Inference</i></b>	= Pengembalian dokumen pada model Probabilistik
<b><i>Teoritical Framework</i></b>	= Teori Kerja
<b><i>Feed Back</i></b>	= Memberi Kembali
<b><i>Keyterm</i></b>	= Kata Kunci
<b><i>Document Matching</i></b>	= Pencocokan Dokumen
<b><i>Similarity measure</i></b>	= Ukuran Kesamaan
<b><i>Term Frequency</i></b>	= Frekuensi Kemunculan Kata dalam sebuah dokumen
<b><i>Inverse Document Frequenc</i></b>	= Kemunculan kata Pada Seluruh dokuemn
<b><i>Direction Difference</i></b>	= Perbedaan Arah
<b><i>Inner Product</i></b>	= Hasil Perkalian
<b><i>Euclidean</i></b>	= Jarak suatu Vektor dangan titik nol
<b><i>Thesaurus</i></b>	= Kamus Sinonim kata
<b><i>Precision</i></b>	= Dokumen relevan dalam dokumen yang ditemukan
<b><i>Recall</i></b>	= Bagian dokumen relevan yang ditemukan
<b><i>Indexing</i></b>	= Memproses Dokumen menjadi kata-kata yang yang akan mewakili dari isi dokumen
<b><i>Pre-procesing</i></b>	= Proses
<b><i>Similarity</i></b>	= ukuran kesamaan
<b><i>Corpus</i></b>	= koleksi dokumen yang diindeks dan dijadikan target pencarian
<b><i>Sampling and reppresentativenes</i></b>	= Sampel dan keterkaitan
<b><i>Finite size</i></b>	= Ukuran yang terbatas
<b><i>Machine-readable form</i></b>	= Bisa dibaca oleh mesin
<b><i>A standard reference</i></b>	= Memiliki standar rukujukan
<b><i>Relevance judgement</i></b>	= sebagai daftar dokumen yang relevan dengan topik-topik yang tersedia
<b><i>Tokenization</i></b>	= Pemisahan rangkaian kata
<b><i>Context Diagram</i></b>	= Gambaran umum dari sistem yang akan dibangun.
<b><i>Database</i></b>	= Basis data yang berisi kumpulan data-data hasil pengamatan.

<b><i>Data Flow Diagram</i></b>	= Menggambarkan suatu sistem yang telah ada atau sistem baru yang akan dikembangkan.
<b><i>Entity Relationship Diagram</i></b>	= Objek data dan hubungan antar diagram.
<b><i>Form</i></b>	= Bentuk dari sebuah tampilan.
<b><i>Implementasi</i></b>	= Pelaksanaan atau penerapan.
<b><i>Informasi</i></b>	= Penerangan, pemberitahuan, kabar atau berita tentang sesuatu.
<b><i>Input</i></b>	= Data yang dimasukkan.
<b><i>Interface</i></b>	= Tampilan antar muka.
<b><i>Output</i></b>	= Data yang dihasilkan.
<b><i>Pre-Processing</i></b>	= Pemrosesan pendahuluan dan pembersihan data merupakan operasi dasar seperti penghapusan noise dilakukan.
<b><i>Prosedur</i></b>	= Tahap kegiatan untuk menyelesaikan suatu aktivitas atau metode langkah demi langkah secara pasti dalam memecahkan suatu masalah.
<b><i>Proses</i></b>	= Runtunan perubahan dalam perkembangan sesuatu.
<b><i>Testing</i></b>	= Pengujian (percobaan) untuk mengetahui tingkat kemampuan atau mengetahui mutunya.
<b><i>User</i></b>	= Pemakai

## DAFTAR LAMPIRAN

Lampiran	Halaman
A. Pengujian Kerelevanan .....	A-1



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Semakin dibutuhkannya informasi membuat perkembangan ilmu pengetahuan yang mendukung sistem temu kembali informasi (STBI) menjadi lebih maju dan berkembang dengan pesat, hal ini terlihat dari banyaknya metode yang ditemukan untuk menyempurnakan metode-metode yang telah ada, salah satunya ialah metode model ruang vektor (*Vector Space Model*) dimana metode ini sangat populer dalam STBI karena metode ini bisa memberikan hasil yang baik serta metode ini juga dapat menampilkan hasil temu balik secara terurut dengan perhitungan aljabar linier (Bunyamin, 2005).

Meskipun STBI telah memiliki peran yang baik akan tetapi ukuran kemiripan yang dimiliki menjadi tidak maksimal jika diterapkan pada teks pendek sebagai kata (*query*) hal ini dikarenakan STBI membandingkan istilah *query* dan dokumen hanya pada tingkat kata atau kalimat dan tidak sampai pada pergeseran *query* yang pendek dimana rata-rata *query* dari pengguna kurang dari dua kata atau hanya dua kata (Imran, dkk., 2009). Sebagai contoh, kata sekolah dimana kata ini umumnya digunakan untuk sebuah tempat pendidikan di Indonesia dan sama halnya dengan kata kampus, madrasa dan pesantren yang juga memiliki makna sebuah tempat atau lembaga pendidikan yang ada, tetapi memiliki perbedaan pada penamaan saja, kasus seperti ini didalam perhitungan kemiripan mengatakan bahwa kata-kata diatas tidak mirip (Bunyamin, 2005).

Untuk mengatasi permasalahan mengenai STBI diatas baik itu mengenai kata yang memiliki makna yang sama maka diusulkan dengan melakukan perluasan *query* (*Query Expansion*), dimana perluasan *query* adalah proses mereformulasikan kembali *query* awal dengan melakukan penambahan beberapa *term* atau kata pada *query* untuk meningkatkan perfoma dalam proses STBI. Dalam konteks *web* mesin pencarian, hal ini termasuk evaluasi *input* pengguna

dan memperluas *query* pencarian untuk mendapatkan dokumen yang cocok dengan *query* (Qiu, 1993).

Akan tetapi didalam penerapan perluasan *query* terdapat permasalahan pemilihan kata berdasarkan kata yang ingin diperluas seperti tidak adanya kata yang ingin dijadikan perluasan atau tidak cocoknya kata yang dijadikan perluasan *query* (imran, dkk., 2009). oleh sebab itu dibuatlah sebuah koleksi kata yang didalamnya terdapat kumpulan kata yang saling bertalian maknanya atau yang dikenal dengan tesaurus, pada saat ini tesaurus telah sering dimasukkan dalam sistem temu kembali informasi untuk mengidentifikasi identik ekspresi bahasa yang secara semantik persis.

Melihat permasalahan yang telah diuraikan dan beberapa penelitian mengenai sistem temu kembali informasi yang sudah pernah dilakukan, maka penulis tertarik untuk penelitian mengenai rancang bangun *query expansion* pada sistem temu kembali informasi (*information retrieval*) dengan model ruang vektor.

## **1.2 Rumusan Masalah**

Berdasarkan dari latar belakang permasalahan, maka dibuat perumusan masalah yang akan dilakukan :

“Bagaimana menerapkan *query expansion* pada sistem temu kembali informasi dengan model ruang vektor ?”

## **1.3. Batasan Masalah**

Adapun yang menjadi batasan masalah dalam laporan tugas akhir ini adalah :

1. Koleksi dokumen yang digunakan adalah tugas akhir dan kerja praktek Mahasiswa jurusan Teknik Informatika Fakultas Sains dan Teknologi UIN Suska Riau.
2. Sistem hanya memproses file Abstrak yang ada pada laporan kerja praktek dan tugas akhir mahasiswa.

3. Untuk *query expansion* akan menggunakan kamus Tesaurus bahasa Indonesia sebagai perluasan *query* yang nantinya digunakan.

## 1.4 Tujuan

Adapun tujuan yang penulis ingin capai didalam penulisan skripsi dan penelitian ini ialah:

1. Terciptanya sebuah sistem temu kembali informasi yang bisa mereformasi *query* dari pengguna yang nantinya akan mengelola data dari Skripsi dan Kerja Praktek dari Mahasiswa Fakultas sains dan Teknologi UIN Suska Riau, agar kelak dapat memberikan kemudahan bagi mahasiswa untuk mendapatkan informasi mengenai Skripsi dan kerja Praktek yang telah ada serta mendapatkan hasil pencarian yang lebih maksimal.
2. Untuk membandingkan sitem temu kembali informasi dengan model ruang vektor tanpa *query expansion* dengan setelah menggunakan *query expansion*.

## 1.5 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini terdiri dari pokok-pokok permasalahan yang dibahas pada masing-masing yang diuraikan menjadi beberapa bagian :

### BAB I PENDAHULUAN

Bab ini membahas tentang gambaran umum isi tugas akhir yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan sistematika penulisan.

### BAB II LANDASAN TEORI

Bab ini menjelaskan tentang teori-teori yang berhubungan dengan penelitian yang terdiri dari penjelasan mengenai konsep sistem temu kembali informasi, model ruang vektor, penerapan *query expansion* pada model ruang vektor.

### BAB III METODOLOGI PENELITIAN

Bab ini membahas langkah-langkah yang dilaksanakan dalam proses penelitian, yaitu pemilihan lokasi penelitian, pengumpulan data, analisa sistem, perancangan sistem, serta pengujian sistem.

### BAB V ANALISA DAN PERANCANGAN SISTEM

Berisikan mengenai pembahasan analisa sistem meliputi analisa data sistem temu kembali informasi, analisa tampilan sistem dan perancangan pada sistem temu kembali informasi, analisa tampilan dari sistem temu kembali informasi, pengkodean, penerapan model ruang vektor, penerapan peluasan *query* sehingga terbentuklah sebuah sistem temu kembali informasi yang utuh.

### BAB VI IMPLEMENTASI

Bab ini akan membahas tentang uji coba sistem yang telah dirancang serta melakukan pengujian *recall* dan *precision* terhadap sistem yang dibuat untuk mengetahui apakah sistem tersebut telah dapat memaksimalkan temu kembali informasi pengelolaan data dan informasi yang ada.

### BAB VII PENUTUP

Pada bab ini akan menjelaskan mengenai kesimpulan dan saran yang nantinya akan didapatkan setelah penelitian ini selesai.

## BAB II

### LANDASAN TEORI

Penyusunan tugas akhir ini membahas mengenai perluasan *query* (*query expansion*) pada temu kembali informasi (*information retrieval*) dengan *model* ruang *vektor*. Sehingga pembahasan teori yang mendukung pelaksanaan tugas akhir ini akan membahas teori mengenai *query expansion* pada temu kembali informasi dengan *model* ruang *vektor* (*vector space model*).

#### 2.1 Sistem Temu Kembali Informasi

Sistem temu kembali informasi (STBI) atau *information retrieval* (IR) digunakan untuk menemukan kembali (*retrieve*) informasi-informasi yang relevan terhadap kebutuhan pengguna dari suatu kumpulan informasi secara otomatis (Mandala, 2002).

##### 2.1.1 Pengertian Sistem Temu Kembali Informasi

Adapun beberapa pendapat para ahli mengenai sistem temu kembali informasi sebagai berikut :

1. Menurut Salton (1989) menyebutkan bahwa STBI suatu proses untuk mengidentifikasi, mengenali dan memanggil dokumen tertentu dalam rangka memberikan jawaban atas permintaan informasi. Dari pendapat Salton diatas dapat disimpulkan bahwa, "Terpanggilnya tidaknya suatu dokumen tergantung dengan kesamaan *query* dengan wakil dokumen".
2. Menurut Lancaster (1998) menyatakan bahwa sistem temu kembali informasi (STBI) adalah suatu proses pencarian dokumen dengan menggunakan istilah-istilah pencarian untuk mendefinisikan dokumen sesuai dengan subjek yang diinginkan.

3. Temu kembali informasi merupakan kegiatan yang bertujuan untuk menyediakan dan memasok informasi bagi pengguna sebagai jawaban atas permintaan berdasarkan kebutuhan pengguna (Basuki 1991).

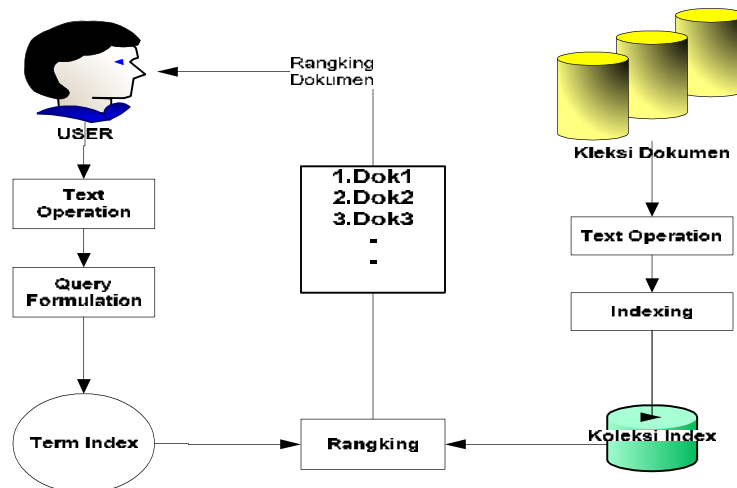
### 2.1.2 Tujuan dan Fungsi dari Sistem Temu Kembali Informasi

Tujuan dari STBI adalah memenuhi kebutuhan informasi pengguna dengan me-*retrieve* semua dokumen yang mungkin relevan, pada waktu yang sama me-*retrieve* sesedikit mungkin dokumen yang tidak relevan. STBI yang baik memungkinkan pengguna menentukan secara cepat dan akurat apakah isi dari dokumen yang diterima memenuhi kebutuhannya (Murad, 2007 ).

Adapun Tujuan dari sistem temu kembali informasi yang ideal (Mandala dan Setiawan 2002) adalah :

1. Menemukan seluruh dokumen yang relevan terhadap suatu *query*.
2. Hanya menemukan dokumen relevan saja, artinya tidak terdapat dokumen yang tidak relevan pada dokumen hasil pencarian.

Sebagai suatu sistem, sistem temu balik informasi memiliki beberapa bagian yang membangun sistem secara keseluruhan. Gambaran bagian-bagian yang terdapat pada suatu sistem temu balik informasi digambarkan pada Gambar 2.1



Gambar 2.1 Bagian-bagian Sistem Temu Balik Informasi (Mandala dan Setiawan, 2002)

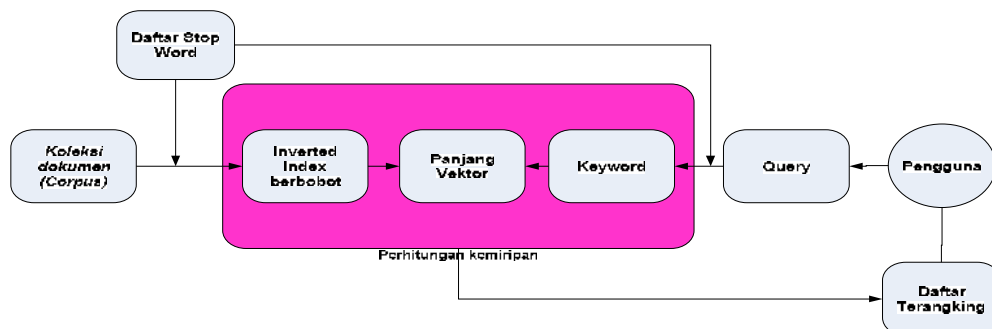
Gambar 2.1 memperlihatkan bahwa terdapat dua buah alur operasi pada sistem temu balik informasi. Alur pertama dimulai dari koleksi dokumen dan alur kedua dimulai dari *query* pengguna. Alur pertama yaitu pemrosesan terhadap koleksi dokumen menjadi basis data indeks tidak tergantung pada alur kedua. Sedangkan alur kedua tergantung dari keberadaan basis data indeks yang dihasilkan pada alur pertama.

Bagian-bagian dari sistem temu balik informasi menurut *Gambar 2.1* meliputi :

1. *Text Operations* (operasi terhadap teks) yang meliputi pemilihan kata-kata dalam *query* maupun dokumen (*term selection*) dalam pentransformasian dokumen atau *query* menjadi *terms index* (indeks dari kata-kata).
2. *Indexing* (pengindeksan), membangun basis data indeks dari koleksi dokumen. Dilakukan terlebih dahulu sebelum pencarian dokumen dilakukan.
3. *Query formulation* (formulasi terhadap *query*) yaitu memberi bobot pada indeks kata-kata *query*.
4. Perangkingan (*ranking*), mencari dokumen-dokumen yang relevan terhadap *query* dan mengurutkan dokumen tersebut berdasarkan kesesuaiannya dengan *query*.

### 2.1.3 Arsitektur Sistem Temu Balik Informasi

Menurut Trunojoyo, arsitektur sistem temu balik informasi dapat digambarkan seperti pada Gambar 2.2 :



Gambar 2.2 Arsitektur Sistem Temu Balik Informasi (Trunojoyo, 2010)

Ada dua hal yang dilakukan sistem pada *Gambar 2.2* diatas, yaitu melakukan *pre-processing* terhadap database dan kemudian menerapkan metode tertentu untuk menghitung kedekatan (relevansi atau *similarity*) antara dokumen di dalam database yang telah dipreproses dengan *query* pengguna.

#### **2.1.3.1 Koleksi Dokumen (Corpus)**

Istilah corpus pada prinsipnya bermakna koleksi dokumen yang diindeks dan dijadikan target pencarian. Suatu *corpus* modern memiliki beberapa karakteristik yakni (McEnery dan Wilson, 2001) :

1. Sampel dan keterkaitan (*Sampling and representativeness*)
2. Ukuran yang terbatas (*Finite size*)
3. Bisa dibaca oleh mesin (*Machine-readable form*)
4. Memiliki standar rujukan (*A standard reference*)

Suatu corpus pengujian sistem temu balik informasi terdiri dari:

1. Koleksi dokumen.
2. Topik-topik, yang dapat digunakan sebagai *query*.
3. Relevance judgement, sebagai daftar dokumen yang relevan dengan topik-topik yang tersedia.

#### **2.1.3.2 Query**

*Query* ialah kata atau ekspresi dari kebutuhan yang dimasukkan pengguna lalu dikonversi sesuai aturan tertentu untuk mengekstrak *term-term* penting yang sejalan dengan *term-term* yang sebelumnya telah diekstrak dari dokumen dan kemudian dihitung relevansi antara *query* dan dokumen berdasarkan pada *term-term* tersebut. Sebagai hasilnya, sistem mengembalikan suatu daftar dokumen terurut (ranking) sesuai nilai kemiripannya dengan ekspresi pengguna (Cios, 2007).



### 2.1.3.3 Inverter *Index* Berbobot

Pembangunan *index* dari koleksi dokumen merupakan tugas pokok pada tahapan preprocessing di dalam sistem temu balik informasi. Kualitas *index* mempengaruhi efektivitas dan efisiensi sistem temu balik informasi. *Index* dokumen adalah himpunan *term* yang menunjukkan isi atau topik yang dikandung oleh dokumen.

*Index* akan membedakan suatu dokumen dari dokumen lain yang berada di dalam koleksi. Ukuran *index* yang kecil dapat memberikan hasil buruk dan mungkin beberapa item yang relevan terabaikan. *Index* yang besar memungkinkan ditemukan banyak dokumen yang relevan tetapi sekaligus dapat menaikkan jumlah dokumen yang tidak relevan dan menurunkan kecepatan pencarian (Trunojoyo, 2010).

Langkah pembangunan inverted *index* (Trunojoyo, 2010) yaitu:

1. Penghapusan tanda baca dan markup pada dokumen Tahap ini menghapus semua tanda baca dan tag markup dari dokumen.

2. Pemisahan rangkaian kata (*tokenization*)

*Tokenization* adalah tugas memisahkan deretan kata di dalam kalimat, paragraf atau halaman menjadi token atau potongan kata tunggal atau *termmed word*. Tahapan ini juga menghilangkan karakter-karakter tertentu seperti tanda baca dan mengubah semua token ke bentuk huruf kecil (*lower case*).

3. Penyaringan (*filtration*)

Pada tahapan ini ditentukan *term* mana yang akan digunakan untuk merepresentasikan dokumen sehingga dapat mendeskripsikan isi dokumen dan membedakan dokumen tersebut dari dokumen lain di dalam koleksi. *Term* yang sering digunakan dianggap sebagai *stop-word* dan dihapus. Penghapusan *stop-word* dari dalam suatu koleksi dokumen pada satu waktu membutuhkan banyak waktu. Solusinya adalah dengan menyusun

suatu pustaka *stop-word* atau *stop-list* dari *term* yang akan dihapus (Manning, 2008).

#### **2.1.3.4 Membandingkan *Query* dengan Dokumen**

Sistem temu balik informasi menerima *query* dari pengguna, kemudian melakukan perangkingan terhadap dokumen pada koleksi berdasarkan kesesuaiannya dengan *query*. Hasil perangkingan yang diberikan kepada pengguna merupakan dokumen yang menurut sistem relevan dengan *query*.

*Model* sistem temu balik informasi menentukan detail sistem temu balik informasi yaitu meliputi representasi dokumen maupun *query*, fungsi pencarian (retrieval function) dan notasi kesesuaian (relevance notation) dokumen terhadap *query*.

#### **2.1.4 *Model* dalam Sistem Temu Balik Informasi**

Di dalam bidang sistem temu balik informasi, dikenal berbagai *model* untuk menilai secara objektif presisi dari suatu pencarian, antara lain *model Boolean* (*Boolean Model*), *model ruang vektor* (*Vector Space Model*) dan *model probabilistik* (*Probabilistic Model*).

##### **2.1.4.1 *Model Boolean***

*Model Boolean* dalam sistem temu kembali merupakan *model* yang paling sederhana. *Model* ini berdasarkan teori himpunan dan aljabar *Boolean*. Dokumen adalah himpunan dari istilah (*term*) dan *query* adalah pernyataan *Boolean* yang ditulis pada *term*. Dokumen diprediksi apakah relevan atau tidak. *Model* ini menggunakan operator *boolean*. Istilah (*term*) dalam sebuah *query* dihubungkan dengan menggunakan operator AND, OR atau NOT.

Kelebihan dari *model boolean* :

1. *Model Boolean* merupakan *model* sederhana yang menggunakan teori dasar himpunan sehingga mudah diimplementasikan.

2. *Model Boolean* dapat diperluas dengan menggunakan *proximity operator* dan *wildcard operator*.
3. Adanya pertimbangan biaya untuk mengubah *software* dan struktur database, terutama pada sistem komersil.

Kelemahan yang dimiliki *Model boolean* :

1. *Model Boolean* tidak bisa melakukan perangkingan didalam menghasilkan pengembalian dokumen dan dokumen yang diambil hanya dokumen yang benar-benar cocok dengan seluruh *query* yang ada pada konteks *query* yang panjang, sehingga ada hasil yang terlalu banyak dan terlalu sedikit, hal ini mempersulit pengguna didalam mengambil keputusan.
2. Kompleksnya pernyataan *Boolean* yang ada. Sehingga dibutuhkan pengetahuan banyak mengenai *query* dengan *boolean* agar pencarian menjadi efisien
3. Tidak bisa menyelesaikan sebagian sesuai (*partial matching*) pada *query*

#### **2.1.4.2 Model Probabilistik**

*Model* probalistik ialah *model* yang melakukan proses pengambilan dokumen sebagai sebuah *probabilistic inference* (Maulan,2011). Karakteristik *model probabilistik* dalam sistem temu kembali adalah sebagai berikut:

1. Melakukan pendugaan page relevansi dengan menggunakan probabilistik
2. Mempunyai *teoritical framework* yang jelas
  - a. Berdasarkan prinsip *statistic*.
  - b. Relevansi dokumen dapat diperbaharui.
  - c. Pengguna bisa melakukan *feed back*.
3. Ide dasar dari *model* Probabilistik:
  - a. *Query* dapat menghasilkan jawaban yang benar
  - b. Menggunakan indeks *term*
  - c. Menggunakan pendugaan awal
  - d. Menggunakan initial hasil

- e. *Feed back* dari *user* dapat memperbaiki probabilitas dari relevansi.

#### 2.1.4.3. *Model Ruang Vektor*

*Model* berikutnya adalah *model* ruang *vektor* yang saat ini sangat populer dalam sistem temu balik informasi. *Model* ini dapat menampilkan hasil temu balik secara terurut (Imran, dkk., 2009). *Model* ruang *vektor* tidak membutuhkan komputasi yang berlebihan sehingga waktu untuk mengeksekusi akan semakin cepat dan lebih efektif (Ramadhany, 2008) sehingga *model* ini lebih baik dari *model boolean*.

Karakteristik yang dimiliki *model* ruang *vektor*

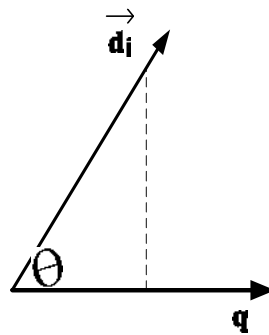
1. *Model vektor* berdasarkan *keyterm*
2. *Model vektor* mendukung *partial matching* (sebagian sesuai) dan penentuan peringkat dokumen
3. Prinsip dasar *model vektor* adalah sebagai berikut :
  - a. Dokumen direpresentasikan dengan menggunakan *vektor keyterm*
  - b. Ruang dimensi ditentukan oleh *keyterms*
  - c. *Query* direpresentasikan dengan menggunakan *vektor keyterm*
  - d. Kesamaan document *keyterm* dihitung berdasarkan jarak *vektor*
4. *Model* ruang *vektor* memerlukan
  - a. Bobot *keyterm* untuk *vektor* dokumen *query*
  - b. Normalisasi *Keyterm* Untuk *vektor* dokument
  - c. Normalisasi *keyterm* untuk *vektor query*
  - d. Perhitungan jarak untuk *vektor* document *keyterm*
5. Kinerja
  - a. Efisien
  - b. Mudah dalam representasi
  - c. Dapat diimplementasikan pada *document matching* dan *partial matching*

Prosedur *model* ruang *vektor* dapat dikelompokkan menjadi tiga tahap yaitu :

- Pengindekan dokumen
- Pembobotan indeks, untuk menghasilkan dokumen yang relevan
- Memberikan peringkat dokumen berdasarkan ukuran kesamaan (*similarity measure*)

Prinsip utamanya adalah *query* diubah menjadi *vektor query* dan dokumen-dokumen di dalam koleksi dokumen diubah menjadi *vektor-vektor* dokumen (Salton, 1975).

Menurut konsep aljabar linier, nilai  $\text{sim}(\vec{d_j}, \vec{q})$  adalah  $\cos \theta$ , seperti pada Gambar 2.3



Gambar 2.3 Besar sudut antara *vektor query* dan *vektor* dokumen

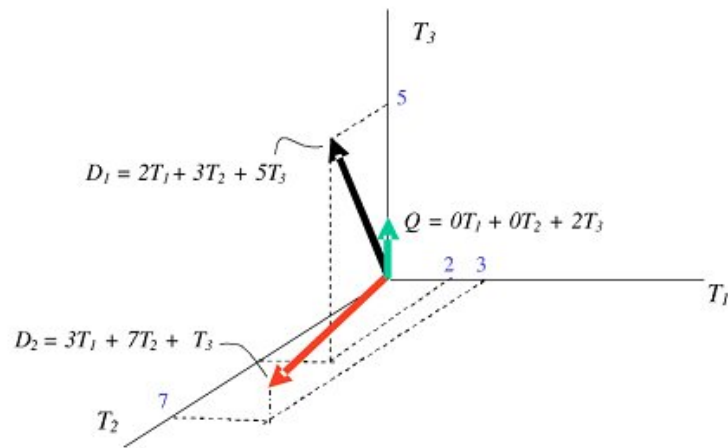
Nilai relevansi terbesar antara dokumen ke- $j$  dan *query* adalah ketika nilai  $\cos \theta$  sama dengan 1 atau nilai  $\theta = 0^\circ$ . Koleksi dokumen direpresentasi pula dalam ruang *vektor* sebagai matriks kata dokumen (*terms documents matrix*). Nilai dari elemen matriks  $W_{ij}$  adalah bobot kata  $i$  dalam dokumen  $j$ . Sebagai contoh terdapat 3 buah kata ( $T_1$ ,  $T_2$  dan  $T_3$ ), 2 buah dokumen ( $D_1$  dan  $D_2$ ) serta sebuah *query*  $Q$ . Masing-masing bernilai :

$$D_1 = 2T_1 + 3T_2 + 5T_3$$

$$D_2 = 3T_1 + 7T_2 + 0T_3$$

$$Q = 0T_1 + 0T_2 + 2T_3$$

Maka representasi grafis dari ketiga *vektor* ini bisa dilihat pada Gambar 2.4 :



Gambar 2.4 Representasi Dokumen dan *Query* pada *Model Ruang Vektor*

#### a. Pembobotan Kata (*Term Weighting*)

Pembobotan kata sangat berpengaruh dalam menentukan kemiripan antara dokumen dengan *query*. Apabila bobot tiap kata dapat ditentukan dengan tepat, diharapkan hasil perhitungan kemiripan teks akan menghasilkan perankingan dokumen yang baik. Keberhasilan dari *model Vektor* ini ditentukan oleh skema pembobotan terhadap suatu *term* baik untuk cakupan lokal maupun global, dan faktor normalisasi (Nicola, 2004). Pembobotan lokal hanya berpedoman pada frekuensi munculnya *term* dalam suatu dokumen dan tidak melihat kemunculan *term* tersebut di dalam dokumen lainnya. Faktor yang memegang peranan penting dalam pembobotan kata, yaitu :

##### 1) *Term Frequency* (tf)

Pendekatan dalam pembobotan lokal yang paling banyak diterapkan adalah *term frequency* (tf). Faktor ini menyatakan banyaknya kemunculan suatu kata dalam suatu dokumen. Semakin sering suatu kata muncul dalam sebuah dokumen, berarti semakin penting kata tersebut. Ada empat cara yang bias digunakan untuk mendapatkan nilai TF (Ramadhany dan Karhendana, 2008):

a) *Raw Tf*

Nilai *Tf* sebuah *term* dihitung berdasarkan kemunculan *term* tersebut dalam dokumen.

b) *Logarithmic Tf*

Dalam memperoleh nilai *Tf*, cara ini menggunakan fungsi logaritmik dalam matematika.  $TF=1+\log (TF)$  ...(2.1)

c) *Binary Tf*

Cara ini, akan menghasilkan nilai *boolean* berdasarkan kemunculan *term* pada dokumen tersebut. Akan bernilai 0 apabila *term* tidak ada pada sebuah dokumen, dan bernilai 1 apabila *term* tersebut ada dalam dokumen. Sehingga banyaknya kemunculan *term* pada sebuah dokumen tidak berpengaruh.

d) *Augmented Tf*

$$TF=0.5+0.5 \times TF_{max} (TF) \quad \dots(2.2)$$

Nilai *Tf* adalah jumlah kemunculan *term* pada sebuah dokumen Nilai  $max(Tf)$  adalah jumlah kemunculan terbanyak *term* pada dokumen yang sama. Dan Perhitungan *Tf* yang akan digunakan dalam implementasi *Information retrieval* System ini adalah *Raw Tf*.

2) *Inverse Document Frequency (idf)*

Pembobotan global digunakan untuk memberikan tekanan terhadap *term* yang mengakibatkan perbedaan dan berdasarkan pada penyebaran dari *term* tertentu di seluruh dokumen. Banyak skema didasarkan pada pertimbangan bahwa semakin jarang suatu *term* muncul di dalam total koleksi maka *term* tersebut menjadi semakin berbeda. Pemanfaatan pembobotan ini dapat menghilangkan kebutuhan *stopwordremoval* karena *stopword* mempunyai bobot global yang sangat kecil. Namun pada prakteknya lebih baik menghilangkan *stopword* di

dalam fase pre-processing sehingga semakin sedikit *term* yang harus ditangani. Pendekatan terhadap pembobotan global mencakup *inverse document frequency* (idf), squared idf, probabilistic idf, GF-idf, entropy. Pendekatan idf merupakan pembobotan yang paling banyak digunakan saat ini. Beberapa aplikasi tidak melibatkan bobot global, hanya memperhatikan tf, yaitu ketika tf sangat kecil atau saat diperlukan penekanan terhadap frekuensi *term* di dalam suatu dokumen. (Polettini, 2004). Bobot global dari suatu *term* i pada pendekatan inverse document frequency (idf) dapat dirumuskan sebagai berikut:

$$Idf = \log \frac{N}{Df} \quad \dots (2.3)$$

Dimana

N = menyatakan jumlah artikel dalam koleksi dokumen

Df = adalah frekuensi dokumen dari *term*.

Log = digunakan untuk memperkecil pengaruhnya relatif terhadap tf

Selanjutnya Bobot *term* dihitung menggunakan :

$$W = tf \times idf \quad \dots (2.4)$$

Dimana :

W = bobot

Tf = *Term Frequency*

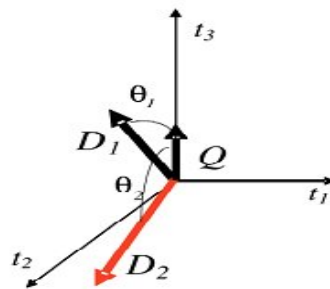
Idf = *inverse document frequency*

## **b. Rumus Relevansi**

Penentuan relevansi dokumen dengan *query* dipandang sebagai pengukuran kesamaan (*similarity measure*) antara *vektor* dokumen dengan *vektor query*. Semakin “sama” suatu *vektor* dokumen dengan *vektor query* maka dokumen dapat dipandang semakin relevan dengan *query*.



Salah satu pengukuran kesesuaian yang baik adalah dengan memperhatikan perbedaan arah (*direction difference*) dari kedua *vektor* tersebut. Perbedaan arah kedua *vektor* dalam geometri dapat dianggap sebagai sudut yang terbentuk oleh kedua *vektor*. Gambar 2.4 mengilustrasikan kesamaan antara dokumen D1 dan D2 dengan *query* Q. Sudut  $\theta_1$  menggambarkan kesamaan dokumen D1 dengan *query* sedangkan sudut  $\theta_2$  menggambarkan kesamaan dokumen D2 dengan *query* sebagai mana pada Gambar 2.5 dibawah ini.



Gambar 2.5 Representasi Grafis Sudut *Vektor* Dokumen dan *Query*

dimana  $Q \cdot D$  adalah hasil perkalian dalam (inner product) kedua *vektor*, sedangkan

$$|Q| = \sqrt{\sum (wq)^2} \text{ dan } |D| = \sqrt{\sum (wd)^2} \quad \dots (2.5)$$

merupakan panjang *vektor* atau jarak Euclidean suatu *vektor* dengan titik nol. Perhitungan kesamaan antara *vektor query* dan *vektor* dokumen dilihat dari sudut yang paling kecil. Sudut yang dibentuk oleh dua buah *vektor* dapat dihitung dengan melakukan perkalian dalam (inner product), sehingga rumus relevansinya adalah:

$$R(Q, D) = \cos \theta = \frac{Q \cdot D}{|Q||D|} \quad \dots (2.6)$$

dimana :

R = Relevansi

Q = Bobot *query*

D = Bobot dokumen

|Q| = Panjang *Vektor query*

|D| = Panjang *Vektor* dokumen

Nilai relevansi (*similarity*) antara *query* dengan dokumen ke-j (Baeza dan Ribeiro, 1999) adalah :

$$\text{similarity}(\vec{d_j}, \vec{q}) = \frac{\vec{d_j} \cdot \vec{q}}{|\vec{d_j}| |\vec{q}|} = \frac{\sum_{i=1}^t (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^t w_{ij}^2 \cdot \sum_{i=1}^t w_{iq}^2}} \dots (2.7)$$

Dimana :

$\vec{d_j}$  = Bobot dokumen j

$\vec{q}$  = Bobot *query*

$|\vec{d_j}|$  = Panjang bobot dokumen j

$|\vec{q}|$  = Panjang bobot *query*

w = Bobot *Term* i pada dokumen j

w = Bobot *term* i pada *query*

t = Teks

### c. Contoh Implementasi

Contoh penggunaan *model* ruang *vektor* dengan data yang digunakan adalah sebagai berikut:

*Query* : “Sistem Informasi”.

Dokumen d1 = Sistem perpustakaan bagian dari sistem informasi.

Dokumen d2 = Informasi dibutuhkan “semua orang”.

Dokumen d3 = Perpustakaan Sekolah.

### **Tahapan preprocessing Dokumen**

Dokumen d1 = Sistem perpustakaan bagian dari sistem informasi

Dokumen d2 = Informasi dibutuhkan semua orang

Dokumen d3 = Perpustakaan Sekolah

### **Tokenisasi (kata diubah ke huruf kecil)**

Dokumen d1 = sistem perpustakaan bagian dari sistem informasi

Dokumen d2 = informasi dibutuhkan semua orang

Dokumen d3 = perpustakaan sekolah

### **Filtering (Penghapusan stop word)**

Daftar stop word : dari, semua

Dokumen d1 = sistem perpustakaan bagian sistem informasi

Dokumen d2 = informasi dibutuhkan orang

Dokumen d3 = perpustakaan sekolah

### ***Indexing dan Pembobotan dokumen***

Pada tahap *indexing*, setelah semua dokumen di proses tiap kata dipisah dan dimasukkan ke dalam tabel *indexing*. Hasil *indexing* dan *weighting* untuk *model* ruang *vektor* dengan data koleksi dokumen di atas dapat dilihat di Tabel 2.1. Dalam koleksi ini, terdapat tiga dokumen, sehingga diperoleh  $N=3$

Tabel 2.1 Hasil Pembobotan Dokumen *Model Ruang Vektor*

Tabel	Term	Id	Tf	Df	Idf	W	D
1	sistem	1	2	2	0.176	0.352	Dok 1= 0.609
2	informasi	1	1	2	0.176	0.176	
3	pustaka	1	1	1	0.477	0.477	
4	bagian	1	1	1	0.477	0.477	
5	informasi	2	1	2	0.176	0.176	Dok 2= 0.487
6	Butuh	2	1	1	0.477	0.477	
7	Orang	2	1	1	0.477	0.477	
8	pustaka	3	1	2	0.176	0.176	Dok 3= 0.258
9	sekolah	3	1	1	0.477	0.477	

Dengan *term* =kata, *id*=nomor dokumen, *TF*=*Term* frekuensi, *Df*=okumen frekuensi, *Idf*=invers dokumen frekuensi, *W*=bobot dan *|D|*=panjang dari tiap dokumen

### Tahapan Preprocessing *Query*

*Query* : “Sistem Informasi”.

Tokenisasi (penghilangan tanda baca)

*Query* q1 = Sistem

*Query* q2 = Informasi

### Tokenisasi

*Query* q1 = sistem

*Query* q2 = informasi

### ***Filtering***

*Query* q1 = sistem

*Query* q2 = informasi

### ***Weighting Query***

Pada tahap pemmerosesan *query* setelah *query* dimasukan pengguna maka akan dilakukan perhitungan bobot dan panjang dari *query* yang ada. Hasil perhitungan bobt dan panjang *query* pada contoh ini dapat dilihat di Tabel 2.2. Dalam koleksi ini, terdapat tiga dokumen yang akan dibandingkan dengan *quey* yang ada, serta kita ketahui bahwa  $N = 3$ .

Tabel 2.2 Hasil pembobotan *query model* ruang vektor

No	<i>Term</i>	Tf	Tf (d1)	Tf (d2)	Tf (d3)	Df	Idf	W	Q
1	Sistem	1	1	0	0	1	0.477	0.477	0.25
2	Informasi	1	1	1	0	2	0.176	0.176	

Dengan *term* =kata, id=nomor dokumen, TF=*Term* ferekuensi, Df=dokumen frekuensi, Idf=invers dokumen frekuensi, W=bobot dan |Q|=panjang dari tiap *query*

### **Perhitungan Relevansi**

Setelah tabel *indexing* di atas (table 2.1) diperoleh, kemudian dilakukan perhitungan terhadap bobot dan panjan *query* yang ada (tabel 2.2). Maka kita akan melakukan perhitungan kemiripan antara *query* dan koleksi dokumen yang ada dengan menggunakan rumus Relevansi *similarity*.

$$R(Q, D) = \cos \theta = \frac{Q \cdot D}{|Q||D|}$$

$$d_j, q = \frac{\vec{d_j} \cdot \vec{q}}{\|\vec{d_j}\| \|\vec{q}\|} = \frac{\sum (d_{j,i} \cdot q_i)}{\sqrt{\sum d_{j,i}^2} \cdot \sqrt{\sum q_i^2}}$$

Dengan mengacu pada hasil tabel 2.1 dan tabel 2.2 maka didapatkan

$$\text{sim}(D1, Q) = \frac{\sum (d_{1,i} \cdot q_i)}{\sqrt{\sum d_{1,i}^2} \cdot \sqrt{\sum q_i^2}}$$

$W_{1,1}$  = Bobot dokumen 1 pada kata 1

$W_{1,2}$  = Bobot *query* 1

$\sum$  = Panjang dokumen 1

$\sum$  = Panjang *query*

$$= \frac{[(1, 0) (1, 0)] + [(0, 1) (0, 1)]}{\sqrt{1^2 + 0^2} \cdot \sqrt{1^2 + 0^2}}$$

$$= \frac{(1, 0) \cdot (1, 0)}{\sqrt{1} \cdot \sqrt{1}}$$

$$= \frac{1}{1}$$

$$= 0.50$$

$$\text{sim}(D1, Q) = 0.50$$

Dengan cara diatas juga diperoleh:  $\text{sim}(D2, Q) = 0,087$  dan  $\text{sim}(D3, Q) = 0$

### Pengurutan Hasil

Dari hasil penghitungan relevansi pada langkah sebelumnya, maka koleksi dokumen tersebut dapat diurutkan dari yang paling relevan sebagai berikut:

Dengan *query* “Sistem Informasi” maka didapat ranking= {d1=1, d2=2, d3=3}

### 2.1.5 Pengujian Sistem Temu Kembali Informasi

keadaan inilah yang digunakan untuk menghitung performansi sistem temu kembali, yaitu *recal* dan *Precision*.

$$Precision = P = tp / (tp + fp) \quad \dots(2.8)$$

$$Recal = R = tp / (tp + fn) \quad \dots(2.9)$$

Denga Keterangan Sebagai Berikut :

P = *Precision*

R = *Recal*

tp = Dokumen yang dikembalikan yang relevan

fp = Dokumen yang dikembalikan yang tidak relevan

fn = Dokumen Yang Tidak Dikembalikan Tetapi Relevan

tn = Dokumen Yang Tidak Dikembalikan Dan Tidak Relevan

Dengan Contoh Asumsi Sebagai Berikut :

Tabel 2.3 Contoh Asumsi dari *Precision* dan *recal*

	Relevan	Tidak Relevan
Ditemukan	(tp)	(fp)
Tidak ditemukan	(fn)	(tn)

## 2.2 Tesaurus

(Sugono, 2008) Kata tesaurus berasal dari kata thesauros, bahasa Yunani, yang bermakna ‘khazanah’.Lambat laun, kata tersebut mengalami perkembangan makna, yakni ‘buku yang dijadikan sumber informasi’. Tesaurus berisi kumpulan kata yang saling bertalian maknanya. Pada dasarnya, tesaurus merupakan sarana untuk mengalihkan gagasan ke dalam sebuah kata, atau sebaliknya. Oleh karena itu, lazimnya tesaurus disusun berdasarkan gagasan atau tema. Namun, untuk memudahkan pengguna dalam pencarian kata, penyusunan tesauruspun berkembang, kini banyak tesaurus yang dikemas berdasarkan abjad. Tesaurus

dibedakan dari kamus. Di dalam kamus dapat dicari informasi tentang makna kata, sedangkan di dalam tesaurus dapat dicari kata yang akan digunakan untuk mengungkapkan gagasan pengguna. Dengan demikian, tesaurus dapat membantu pengguna dalam mengungkapkan atau mengekspresikan gagasan sesuai dengan apa yang dimaksud. Misalnya, pencarian kata lain untuk kata hewan, pengguna tesaurus dapat mencarinya pada lema hewan. “hewan n binatang, dabit, fauna, sato, satwa” Sederet kata yang terdapat pada lema hewan tersebut menunjukkan bahwa kata tersebut bersinonim sehingga dapat saling menggantikan sesuai dengan konteksnya. Tesaurus berguna dalam pengajaran bahasa sehingga dapat dimanfaatkan oleh pengajar dan pelajar.

### **2.3. *Query Expansion***

Ukuran kemiripan teks standard, yaitu kemiripan cosinus di dalam ruang *vektor* seperti di atas dapat memberikan hasil tidak memuaskan jika diterapkan pada teks sangat pendek, seperti pada perhitungan kemiripan antar *query*. kata “sekolah” dimana kata ini umumnya digunakan untuk sebuah tempat pendidikan di Indonesia dan sama halnya dengan kata kampus, madrasa dan pesantren yang juga memiliki makna sebuah tempat atau lembaga pendidikan yang ada, tetapi memiliki perbedaan pada penamaan saja, Perhitungan kemiripan standard mengatakan bahwa kata-kata tersebut tidak memiliki kemiripan (Metzler,dkk., 2007) oleh sebab itulah dibutuhkan *query expansion* untuk merepresentasikan kebutuhan pengguna dalam temukembali informasi.

*Query Expansion* atau perluasan *query* adalah proses me-reformulasikan kembali *query* awal dengan melakukan penambahan beberapa *term* atau kata pada *query* untuk meningkatkan perfoma dalam proses *information retrieval*. Dalam konteks web search engine, hal ini termasuk evaluasi input *user* dan memperluas *query* pencarian untuk mendapatkan dokumen yang cocok dengan *query* (Qiu, 1993).

Proses perluasan *query* dalam sistem ini akan dibangun korpus perluasan kata untuk pengguna terlebih dahulu, dengan mengumpulkan secara keseluruhan



kata untuk membantu pengguna dalam merumuskan kembali *query* yang ada dalam hal ini akan dilakukan dengan menggunakan sinonim kata dari thesaurus yang ada.

Untuk aturan *Query Expansion* akan menggunakan aturan pembobotan *Query Expansion* (harjono, 2005):

$$W_{qt} = w \times \text{adj} \quad \dots (2.10)$$

Dimana :

$W$  = bobot

$W_{qt}$  = bobot *term* dalam *query* bila dilakukan *query expansion*

$\text{adj}$  = faktor penyesuaian

dimana kata asli diberi faktor penyesuaian sebesar 1 dan kata perluasan diberikan factor penyesuaian 0.5.

Berikut ini adalah simulasi penerapan *Query Expansion* yang dilakukan dengan melakukan perhitungan *model* ruang *Vektor* :

Dimana

$N=4$

dok 1 = analisa jaringan komputer pada kantor rektorat kampus UIN suska Riau

dok 2= sistem Pembayaran uang bulanan (sekolah menengah kejuruan 2 Pekanbaru)

dok 3 = optimasi website uin suska riau dengan metode search engine optimization di pusat komputer uin suska riau

dok 4 =sistem informasi kesiswaan madrasah aliyah darul hikmah (pondok pesantren dar el hikmah)

*Query* = Sekolah

Tabel 2.4 Pengindekan Dokumen serta perhitungan bobot dokumen

				$\text{idf} = \log(N/\text{df})$	$W = \text{Tf} * \text{idf}$		
<i>Term</i>	<i>Id</i>	<i>Tf</i>	<i>Df</i>	<i>Idf</i>	<i>W</i>	$W^2$	$ D $
Analisa	1	1	1	0,60206	0,60206	0,36248	1,47473977
Jarin	1	1	1	0,60206	0,60206	0,36248	

Tabel 2.4 Lanjutan

<i>Term</i>	<i>Id</i>	<i>Tf</i>	<i>Df</i>	<i>Idf</i>	<i>W</i>	<i>W<sup>2</sup></i>	<b> D </b>
Komputer	1	1	2	0,30103	0,30103	0,09062	
Kantor	1	1	1	0,60206	0,60206	0,36248	
Rektorat	1	1	1	0,60206	0,60206	0,36248	
Kampus	1	1	1	0,60206	0,60206	0,36248	
Uin	1	1	2	0,30103	0,30103	0,09062	
Suska	1	1	2	0,30103	0,30103	0,09062	
Riau	1	1	2	0,30103	0,30103	0,09062	
sistem	2	1	2	0,30103	0,30103	0,09062	1,72928567
Bayar	2	1	1	0,60206	0,60206	0,36248	
Uang	2	1	1	0,60206	0,60206	0,36248	
Bulan	2	1	1	0,60206	0,60206	0,36248	
Sekolah	2	1	1	0,60206	0,60206	0,36248	
Tengah	2	1	1	0,60206	0,60206	0,36248	
Juru	2	1	1	0,60206	0,60206	0,36248	
2	2	1	1	0,60206	0,60206	0,36248	
Pekanbaru	2	1	1	0,60206	0,60206	0,36248	
Optimasi	3	1	1	0,60206	0,60206	0,36248	1.92753246
Website	3	1	1	0,60206	0,60206	0,36248	
Uin	3	2	2	0,60206	1,20412	1,44990	
Suska	3	2	2	0,60206	1,20412	1,44990	
Riau	3	2	2	0,60206	1,20412	1,44990	
Metode	3	1	1	0,60206	0,60206	0,36248	
Search	3	1	1	0,60206	0,60206	0,36248	
Engine	3	1	1	0,60206	0,60206	0,36248	
Optimization	3	1	1	0,60206	0,60206	0,36248	
Pusat	3	1	1	0,60206	0,60206	0,36248	
Komputer	3	1	2	0,60206	0,60206	0,36248	2,27272663
sistem	4	1	2	0,60206	0,60206	0,36248	
Informasi	4	1	1	0,60206	0,60206	0,36248	
Siswa	4	1	1	0,60206	0,60206	0,36248	
Madrasah	4	1	1	0,60206	0,60206	0,36248	
Aliyah	4	1	1	0,60206	0,60206	0,36248	
Darel	4	2	1	0,60206	0,60206	0,36248	
Hikmah	4	2	1	0,60206	1,20412	1,44990	
Pondok	4	1	1	0,60206	0,60206	0,36248	
Pesantrean	4	1	1	0,60206	0,60206	0,36248	

Dengan term =kata, id=nomor dokumen, TF=Term frekuensi, Df=okumen frekuensi, Idf=invers dokumen frekuensi, W=bobot dan  $|D|$ =panjang dari tiap dokumen

Tabel 2.5 perhitungan bobot query

Perhitungan Bobot <i>Query</i>			$idf = \log(N/df)$	$W = Tf * idf$		
Term	Df	Tf	Idf	W	$W^2$	$ Q $
Sekolah	1	1	0,301029996	0,301029996	0,090619	0,301029996

Dengan term =kata, id=nomor dokumen, TF=Term frekuensi, Df=okumen frekuensi, Idf=invers dokumen frekuensi, W=bobot dan  $\sum$ =panjang dari tiap dokumen

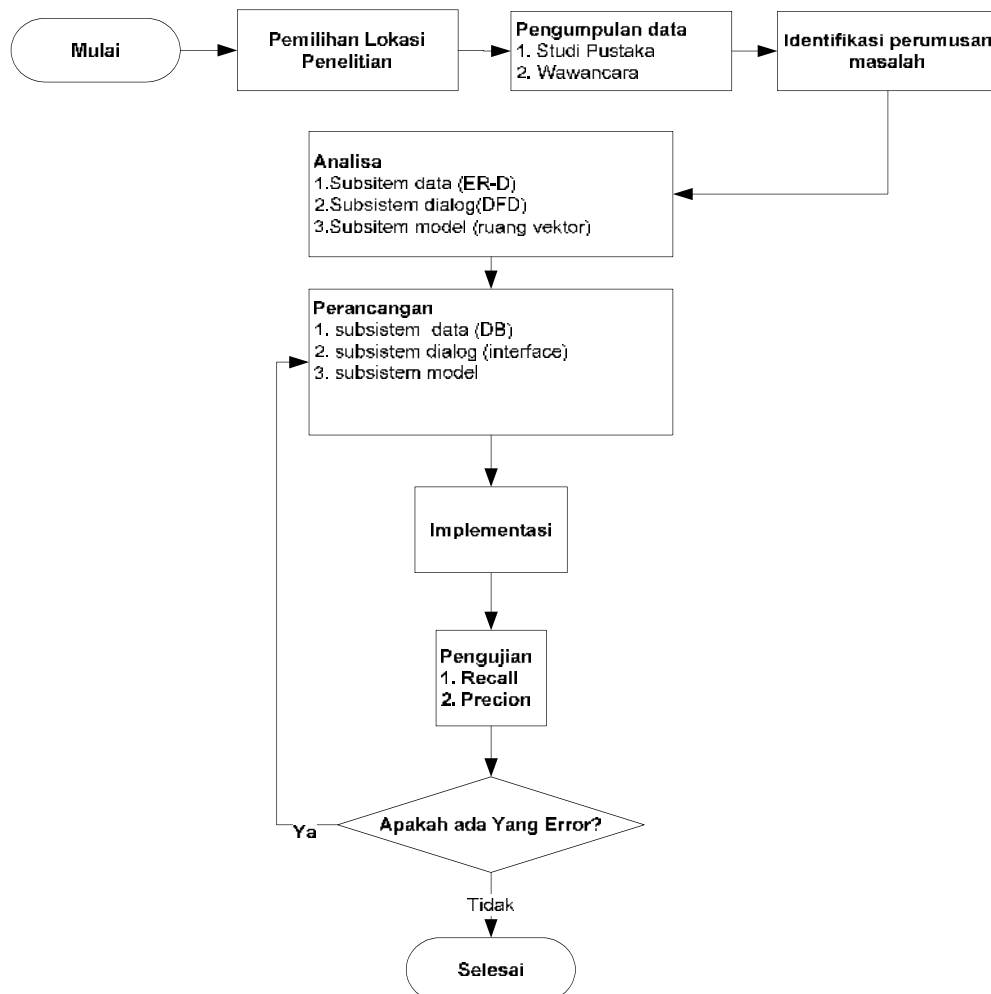
Table 2.6 Perhitungan Simmilarity (dj,q) *Vektor*

<b><i>SIMILARITY (d<sub>i</sub>,q) VEKTOR</i></b>		
dok id	nilai <i>Similarity</i>	Rangking
dok1	0	0
dok2	0,348155312	1
dok3	0	0

## BAB III

### METODOLOGI PENELITIAN

Metodologi penelitian merupakan sistematika tahapan yang akan dilaksanakan selama pengerjaan tugas akhir ini . berikut ini adalah tahapan yang dilalui dalam pengerjaan tugas akhir ini yang digambarkan dalam bentuk *flowchart* sebagai mana pada Gambar 3.1:



Gambar 3.1 *Flowchart* Penyusunan Tugas Akhir

### **3.1 Pemilihan Lokasi Penelitian**

Penelitian akan dilakukan di perpustakaan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau dengan pertimbangan bahwa hasil penelitian akan sistem temu kembali informasi bisa digunakan oleh mahasiswa yang akan mencari informasi mengenai Tugas Akhir dan kerja Praktek

Waktu penelitian berlangsung dari bulan Juni 2012 sampai dengan Agustus 2012 atau selama kurang lebih 2 bulan.

### **3.2 Pengumpulan Data**

Ada dua metode yang penulis lakukan untuk memperoleh informasi atau pengumpulan data pada penelitian ini yaitu metode studi pustaka.

Studi Pustaka berfungsi untuk mendukung penelitian yang akan dilaksanakan. Pengumpulan teori-teori yang mendukung dalam penelitian ini merupakan kegiatan dalam studi pustaka. Sumber yang digunakan dapat berupa buku, jurnal dan tulisan penelitian yang berhubungan dengan yang penulis lakukan sekarang.

### **3.3 Identifikasi dan Perumusan Masalah**

Dengan memanfaatkan informasi hasil studi pustaka dan wawancara yang telah dilakukan, maka dilakukan tahap berikutnya yaitu mengidentifikasi masalah. Pada tugas akhir ini masalah yang akan diidentifikasi adalah bagaimana mereformulasi *query* untuk merepresentasikan kebutuhan pengguna akan informasi yang ada dengan menerapkan *query expansion*. Berdasarkan masalah yang telah diidentifikasi, maka dirumuskanlah suatu masalah, yaitu bagaimana menerapkan *query expansion* pada sistem temu kembali informasi (*information retrieval*) dengan model ruang vektor.

### 3.4 Analisa Sistem

Analisa merupakan metode yang dilakukan setelah merumuskan masalah yang ada, Analisa berarti metode yang khusus untuk menganalisis masalah dibangunnya sistem dan hasil akhir yang ingin dicapai dari pembuatan sistem. Analisa utama yang akan dilakukan adalah :

1. Analisa subsistem data, dimana didalam tahap ini kita akan menganalisa mengenai apa saja kebutuhan data yang akan digunakan didalam membangun penelitian ini.
2. Analisa subsistem dialog, pada tahap ini kita akan menganalisa bagaimana sebuah tampilan dari sistem temu kembali informasi yang akan dibangun, sehingga kelak terbentuklah sebuah tampilan yang *user friendly* yang akan membuat pengguna nyaman dan mudah didalam menggunakannya
3. Analisa subsistem model

Dalam proses ini terdapat beberapa tahapan dalam membangun sistem temu balik informasi yaitu :

- a. Melakukan analisa pemisahan rangkaian kata atau *tokenization*.
- b. Melakukan analisa penyaringan kata (*filtration*) berdasarkan daftar *stop list*.
- c. Analisa bobot terhadap istilah (*weighting*).
- d. Melakukan Normalisasi dokumen
- e. Analisa model ruang vektor yang akan digunakan, prinsip utama dari model ini adalah mengubah *query* yang diinputkan pengguna menjadi vektor *query* dan dokumen pada koleksi dokumen (*corpus*) menjadi vektor dokumen. Semakin “mirip” suatu vektor dokumen dengan vektor *query* maka dokumen dapat dipandang semakin relevan dengan *query*.
- f. Analisa penerapan *Query Expansion*, pada tahap ini kita akan menganalisa bagaimana perluasan sebuah *query* nantinya yang akan diterapkan pada sistem.

### 3.5 Perancangan Sistem

Perancangan berarti metode yang khusus digunakan untuk merancang sistem yang telah dianalisa dengan tujuan untuk memberikan kemudahan dan menyederhanakan suatu proses atau jalannya aliran data, perancangan terhadap model, dan merancang rancang bangun sistem. Adapun rancangan utama sistem yaitu :

1. Perancangan subsistem data, dimana didalam tahap ini kita akan merancang *database* dari sistem, prosesnya meliputi :
  - a. Konversi file untuk diindex kedalam format Xml, dimana nantinya file Xml yang didalamnya terdapat kumpulan kata akan dibaca langsung oleh sistem .
  - b. Melakukan indexing data ke dalam *database*.
2. Perancangan subsistem dialog, pada tahap ini kita akan merancang sebuah tampilan dari sistem temu kembali informasi yang akan dibangun, sehingga kelak terbentuklah sebuah tampilan yang *userfriendly* yang akan membuat pengguna nyaman dan mudah didalam menggunakannya
3. Perancangan subsistem model

Dalam proses ini terdapat beberapa tahapan dalam membangun sistem temu balik informasi yaitu :

  - a. Perancangan model ruang vektor pada sistem temu kembali informasi.
  - b. Penerapan *Query Expansion*

### 3.6 Implementasi

Pada proses implementasi ini akan dilakukan pembuatan modul yang telah dirancang dan dianalisa selanjutnya diimplementasikan pada bahasa pemrograman dan dilakukan pengujian untuk mengetahui tingkat keberhasilan aplikasi yang telah ada. Berikut adalah spesifikasi lingkungan implementasi perangkat keras dan perangkat lunak :

a. Perangkat Keras

1. Processor : Intel Core 2 Duo 2.00 GHz
2. Memory : 1 GB
3. Harddisk : 500 GB

b. Perangkat Lunak

1. Sistem Operasi : Windows 7 Ultimate
2. Browser : Google Chrome
3. Bahasa Pemrograman : Php
4. DBMS : MySQL

### 3.7 Pengujian Sistem

Pengujian merupakan tahapan dimana sistem akan dijalankan. Tahap pengujian diperlukan sebagai ukuran bahwa sistem dapat dijalankan sesuai dengan tujuan. Pengujian sistem temu balik informasi dalam tugas akhir ini dilakukan dengan cara mengukur kualitas dari informasi yang dikembalikan sistem berdasarkan *query* yang diinputkan pengguna. Ukuran yang digunakan untuk mengukur kualitas sistem temu balik informasi ini menggunakan *precision* dan *recall*.

### 3.8 Kesimpulan dan Saran

Tahapan kesimpulan dan saran merupakan akhir dari penelitian tugas akhir. Tahapan kesimpulan membahas hasil evaluasi dari seluruh kegiatan yang dilakukan dalam melakukan penelitian terhadap pembuatan Sistem Temu Balik informasi dengan model ruang vektor yang telah dilengkapi dengan *query expansion* serta memberikan saran-saran untuk menyempurnakan dan mengembangkan penelitian sistem temu balik informasi selanjutnya.



## **BAB IV**

### **ANALISA DAN PERANCANGAN**

Pada perancangan sistem temu kembali informasi, tahap analisa memegang peranan yang penting dalam pelaksanaannya untuk membuat rincian sistem baru yakni berupa langkah pemahaman persoalan sebelum mengambil tindakan atau keputusan penyelesaian hasil utama. Sedangkan tahap perancangan sistem adalah membuat rincian sistem hasil dari analisa menjadi bentuk perancangan, karena diharapkan nantinya sistem akan lebih mudah untuk dimengerti oleh pengguna.

#### **4.1. Analisa Sistem Lama**

Sistem temu kembali informasi telah banyak dikembangkan, salah satunya dengan menggunakan Model ruang vektor, dimana pada saat ini model ini menjadi salahsatu model yang cukup populer didalam sistem pencarian informasi, karena dinilai cukup efektif dan mampu memberikan hasil yang memadai didalam proses temu kembali informasi, akan tetapi model ini menjadi tidak maksimal ketika digunakan pada teks-teks pendek.

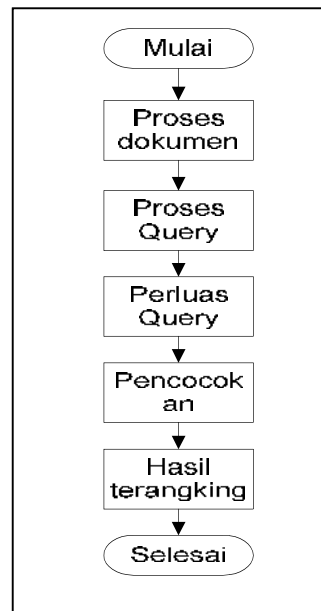
Maka dari itu diusulkanlah penambahan fitur pada sebuah sistem temukembali yang telah menggunakan model ruang vektor yang ada yaitu dengan melakukan penambahan perluasan *query* yang ada, diharapkan dengan adanya penambahan fitur ini bisa lebih memaksimalkan pencarian data yang ada, terutama pada teks-teks pendek.

#### **4.2. Analisa Pengembangan Sistem**

Secara garis besar, ada beberapa tahapan yang ditangani oleh sistem temu kembali informasi dengan model ruang vektor yaitu :

- a. Melakukan proses terhadap dokumen atau yang dikenal dengan pembentukan korpus diantaranya pengindekan, pembobotan dan normalisasi vektor dokumen.
- b. Melakukan preproses terhadap *query* pengguna, dimana nantinya *query* yang diinputkan pengguna akan diproses oleh sistem seperti halnya dengan dokumen, pada *query* ini juga akan dihitung bobot dan dilakukan normalisasi vektor *query*.
- c. *Matching* atau pencocokan antara *query* dan dokumen yang ada, pada proses ini juga akan Menerapkan metode ruang vektor untuk menghitung kedekatan (*relevansi/similarity*) antara dokumen dan *query* pengguna tersebut.
- d. Perengkingan hasil dari pencarian.

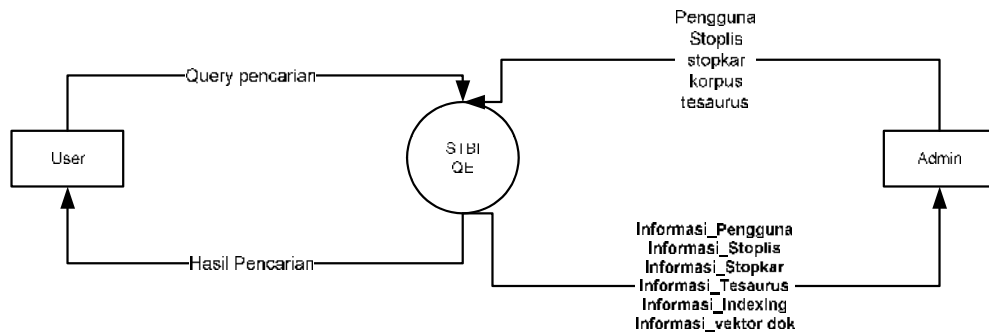
Empat tahap diatas adalah gambaran umum dari sistem temu kembali yang telah ada, dan untuk penelitian kalai ini, tahapan diatas akan dilakukan penambahan pada proses “b”, dimana nantinya *query* akan diperluas dengan menggunakan kamus tesaurus, yang diharapkan bisa menambah relevansi dari temu kembalif yang ada.



Gambar 4.1 Menunjukkan *flowchart* sistem secara umum setelah diterapkan *Query expansion*.

#### 4.2.1. Analisa subsistem dialog

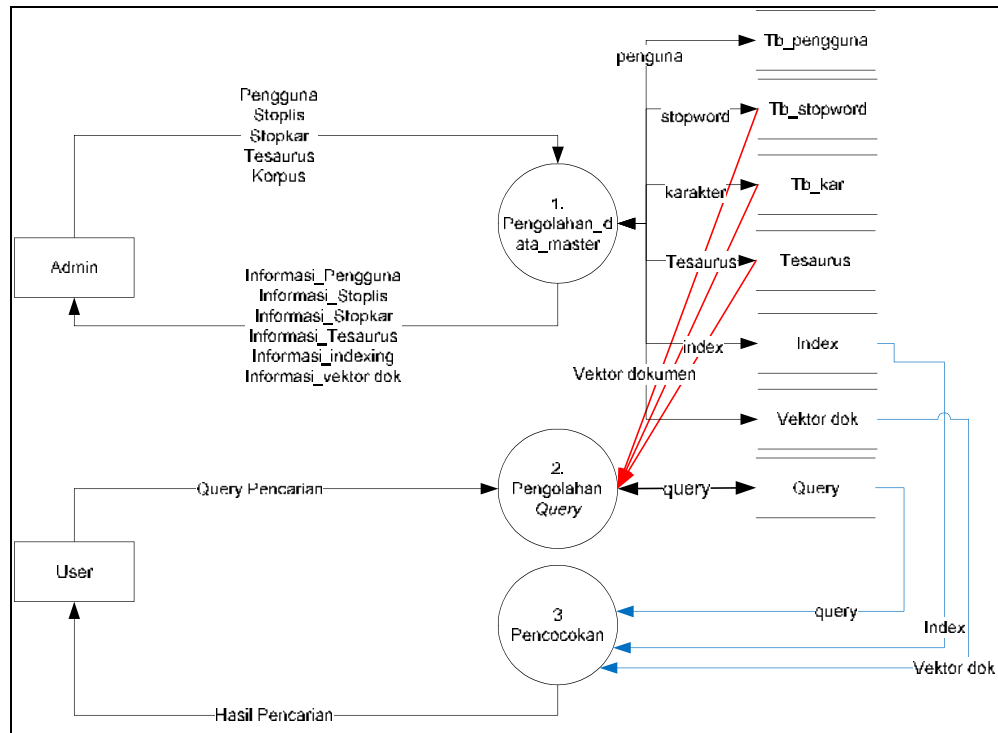
Analisa subsistem dialog dijelaskan dengan menggunakan DFD (data flow diagram), selengkapnya bisa dilihat pada Gambar 4.2 Konteks diagram STBI QE, Tabel 4.2 keterangan konteks diagram, Gambar 4.3 Dfd Lvl 1 dan Tabel 4.2 keterangan Dfd Lvl 1. Gambar 4.4 untuk Dfd Lvl 2 Proses 1 dan Tabel 4.3 keterangan Dfd Lvl 2 Proses 1.



Gambar 4.2 Konteks Diagram STBI QE

Tabel 4.1 Keterangan Context diagram

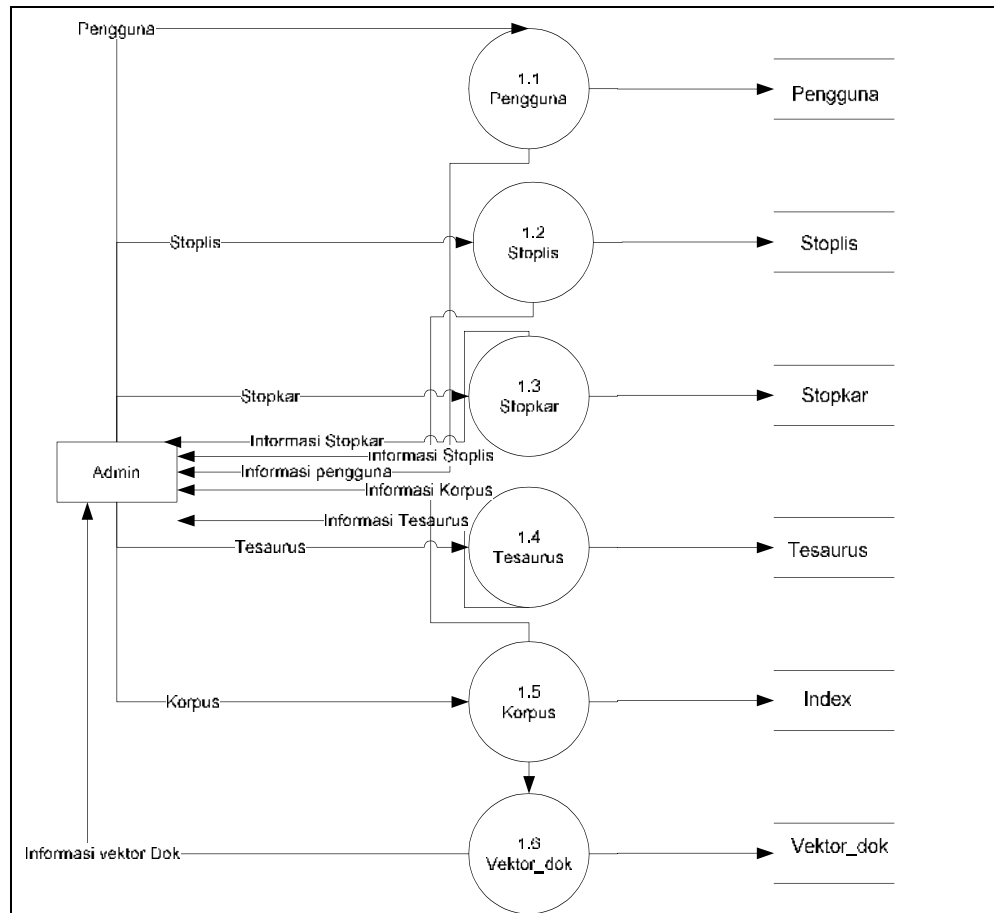
No	Jenis	Nama	Deskripsi
1	Entitas	User	Pengguna dari STBI QE
		Admin	Pengelola sistem
2	Aliran data	Query Pencarian	Kata/kalimat yang dimasuka User
		Hasil Pencarian	Hasil Pencarian yang akan ditampilkan oleh sistem
		Stoplis	Berisikan kata-kata <i>Stopword</i>
		Stopkar	Berisikan Karakter-karakter Penulisan
		Korpus	Koleksi dokumen Pencarian
		Indexing	Hasil pengindekan dokumen
		Vektor dokumen	Berisikan ukuran panjang dari tiap isi dokumen
		Tesaaurus	Berisikan kompul kata dari kamus thesaurus
		Vektor dokumen	Berisikan ukuran panjang dari tiap isi dokumen
		Tesaaurus	Berisikan kompulan kata dari kamus thesaurus
		Informasi Tesaaurus	Berisikan infromasi thesaurus
		Informasi Pengguna	Berisikan informasi data pengguna
		Informasi Stoplis	Berisikan infromasi stoplis
		Informasi Stopkar	Berisikan informasi stopkar



Gambar 4.3 DFD Lvl 1

Tabel 4.2 Keterangan Dfd Level 1

No	Jenis	Nama	Deskripsi
1	Proses	Pengolahan data master	Mengolah seluruh data master pada sistem
		Pengolahan <i>Query</i>	Proses Pengolahan <i>Query</i> dari pengguna
		Pencocokan	Proses Pencocokan antara <i>query</i> dan <i>index</i> yang ada
2	Aliran data	<i>Query</i> Pencarian	Kata/kalimat yang dimasuka User
		Hasil Pencarian	Hasil Pencarian yang akan ditampilkan oleh sistem
		Stoplis	Berisikan kata-kata <i>Stopword</i>
		Stopkar	Berisikan Karakter-karakter Penulisan
		Korpus	Koleksi dokumen Pencarian
		Vektor dokumen	Berisikan ukuran panjang dari tiap isi dokumen
		Tesaurus	Berisikan kompulan kata dari kamus tesaurus
		Informasi Tesaurus	Berisikan infromasi thesaurus
		Informasi Pengguna	Berisikan informasi data pengguna
		Informasi Stoplis	Berisikan infromasi stoplis
		Informasi Stopkar	Berisikan informasi stopkar
		Informasi Korpus	Berisikan informasi korpus



Gambar 4.4 Dfd Levle 2 Proses 1

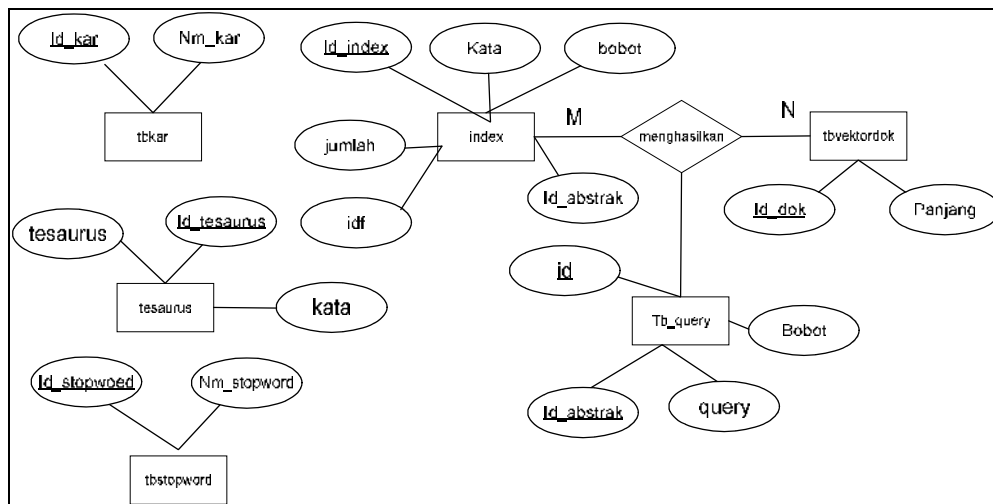
Tabel 4.3 Keterangan DFD Level 2 Proses 1

No	Jenis	Nama	Deskripsi
1	Proses	Pengguna	Proses Pengolahan data pengguna
		Stoplis	Proses Pengolahan data Stoplis
		Stopkar	Proses pengolahan data Stopkar
		Tesaurus	Proses pengolahan data thesaurus
		Korpus	Proses pengolahan data korpus
		Vektor dok	Proses pengolahan vektor dokumen
2	Aliran data	Stoplis	Berisikan kata-kata <i>Stopword</i>
		Stopkar	Berisikan Karakter-karakter Penulisan
		Korpus	Koleksi dokumen Pencarian
		Vektor dokumen	Berisikan ukuran panjang dari tiap isi dokumen
		Tesaurus	Berisikan kumpulan kata dari tesaurus
		Kata Dasar	Berisikan Kata dasar kamus Bahasa Indonesia
		Informasi Tesaurus	Berisikan infromasi thesaurus
		Informasi Pengguna	Berisikan informasi data pengguna

	Informasi Stoplis	Berisikan infromasi stoplis
	Informasi Stopkar	Berisikan informasi stopkar
	Informasi Korpus	Berisikan informasi korpus

#### 4.2.2. Analisa Subsistem Data

Pada analisa data sistem akan dijelaskan mengenai *Entity Relation Diagram* (ERD) yang dapat dilihat pada Gambar 4.4 dan keterangannya dapat dilihat pada Tabel 4.3



Gambar 4.5 ERD

Tabel 4.4 ERD Sistem

N o	Jenis	Nama	Deskripsi	Atribut	Primary Key
1	E	Tbkar	Menyimpan data karakter dan symbol penulisan	Id_kar Nm_kar	Id_kar
	T	Thesaurus	Menyimpan data kata kamus thesaurus	id_tesaurus kata thesaurus	id_tesaurus
	I	tbstopword	Menyimpan kumpulan kata yang dianggap tidak memiliki peran penting pada hasil pencarian, atau memiliki jumlah	Id_stopword Nm_stopword	Id_stopword

			kemunculan yang sangat sering		
		Tbvektordok	Menyimpan nilai vektor dari setiap	id_dok panjang	id_dok
		<i>Index</i>	Menyimpan hasil pengindekan dari dokumen korpus yang ada	Id_index Kata Bobot Jumlah Idf Id_dokumen	Id_index
		Tbquery	Menyimpan hasil <i>query</i> pencarian pengguna, dan bersifat sementara.	Id <i>Query</i> Id_abstrak Bobot	Id

#### 4.2.3. Analisa model

Analisa model *Query* expansion pada sistem temu kembali informasi ini akan membahas mengenai pembentukan korpus, proses *query*, perluasan *query*, pencocokan antara *query* dengan indenk (*similarity*) dengan menggunakan model ruang vektor dan yang terakhir perangkikan.

##### 4.2.3.1. Pembentukan korpus

Dalam pembentukan korpus ini, akan melalui beberapa tahap, dari mulai persiapan file korpus dokumen, lalu dilakukan pengindekan hingga didapat sebuah tabel *index* yang didalamnya telah terdapat panjang dari setiap dokumen yang ada.

##### a. File korpus

Dokumen yang akan dijadikan korpus adalah dokumen abstrak dari kerja praktek dan Tugas akhir mahasiswa teknik informatika yang nantinya akan dirubah kedalam format .xml, dengan urutan tag xml sebagai berikut :

```

<?xml version="1.0" encoding="UTF-8"?> header standard XML
<doc> root
<docno>berisikan no dokumen </docno> child.1
<title> Judul dari TA/KP</title> child.2
<author> Nama Penulis dan Nim </author> child.3
<date> tanggal keluar TA/KP</date> child.4
<text> ABSTRAK DARAI TA/KP</text> child.5
</doc> end root

```

## b. pengindekan

Setelah dokumen korpus selesai di buat, maka proses selanjutnya iyalah pengindekan isi file Xml ke dalam *database* dan hal itu melalui beberapa tahap :

### 1.tokenisasi

#### Pemecahan Kata

Contoh :

Berikut ini adalah contoh dokumen yang akan disnimpan kedalam koleksi dokumen :

dok 1 = analisa jaringan komputer pada kantor rektorat kampus UIN suska Riau

dok 2= sistem Pembayaran uang bulanan (sekolah menengah kejuruan 2 Pekanbaru)

dok 3 = optimasi website uin suska riau dengan metode search engine optimization di pusat komputer uin suska riau

dok 4 =sistem informasi kesiswaan madrasah aliyah darul hikmah (pondok pesantren dar el hikmah)

#### **Input :**

analisa jaringan komputer pada kantor rektorat kampus UIN suska Riau

sistem Pembayaran uang bulanan (sekolah menengah kejuruan 2 Pekanbaru)

optimasi website uin suska riau dengan metode search engine optimization di pusat komputer uin suska riau

sistem informasi kesiswaan madrasah aliyah derel hikmah (pondok pesantren darel hikmah)



**Output :**

analisa jaringan komputer pada kantor rektorat kampus UIN suska Riau  
sistem Pembayaran uang bulanan ((sekolah menengah kejuruan 2 Pekanbaru))  
optimasi website uin suska riau dengan metode search engine optimization di pusat  
komputer uin suska riau  
sistem informasi kesiswaan madrasah aliyah darul hikmah ((pondok pesantren  
darel hikmah))

**Merubah ke huruf kecil**

Setelah dilakukan pemecahan, sehingga menjadi kata dan kata,  
kemudian dilakukan proses perubahan setiap huruf besar kedalam huruf  
kecil.

**Output :**

analisa jaringan komputer pada kantor rektorat kampus uin suska riau  
sistem pembayaran uang bulanan ((sekolah menengah kejuruan 2 pekanbaru))  
optimasi website uin suska riau dengan metode search engine optimization di pusat  
komputer uin suska riau  
sistem informasi kesiswaan madrasah aliyah darul hikmah ((pondok pesantren  
darel hikmah))

Keterangan :

UIN = uin      Riau = riau      Pembayaran = pembayaran

Pekanbaru = pekanbaru

**Penyaringan (*filtering*)**

Pada tahapan akan dilakukan penghapusan stoplis dan karakter dan  
diganti dengan spasi “ ”.

**Output penghapusan stopwords :**

analisa jaringan komputer pada kantor rektorat kampus uin suska riau  
sistem pembayaran uang bulanan sekolah menengah kejuruan 2 pekanbaru  
optimasi website uin suska riau metode search engine optimization pusat komputer  
uin suska riau  
sistem informasi kesiswaan madrasah aliyah darul hikmah ((pondok pesantren  
darel hikmah))

Di,dengan, (, ) .

### **Pembobotan**

Setelah proses penghapusan atau penyarian selesai, maka kata telah bisa disimpan kedalam *database* :

Tabel 4.5 Penyimpanan kata kedalam *database*

Analisa	Metode
Jaring	Search
Komputer	Engine
Kantor	Optimization
Rektorat	Pusat
Kampus	Komputer
Uin	sistem
Suska	Informasi
Riau	Siswaan
sistem	Madrasah
Bayar	Aliyah
Uang	Darel
Bulan	Hikmah
Sekolah	Pondok
Tengah	Pesantrean
Juru	
2	
Pekanbaru	
Optimasi	
Website	
Uin	
Suska	
Riau	

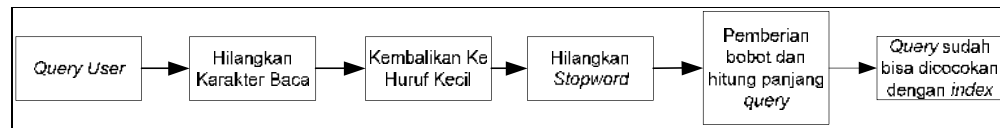
Kemudikan akan dilakukan perhitungan bobot dari setiap kata yang ada, untuk pembobotan lokal (TF) dimana untuk judul dari abstrak yang ada kan diberikan nilai 1 pada setiap kali kemunculannya akan tetapi untuk bagian isi dari abstrak akan diberikan bobot 0.4 pada setiap kemunculannya, untuk pembobotan Global (Idf) dilakukan dama dengan persamaan (2.3) dan untuk perhitungan panjang vektor dokumen dilakukan perhitungan berdasarkan persamaan (2.5)

Tabel 4.6 Perhitungan Bobot kata

				idf=Log(N/df)	$W=Tf*Idf$	
<b>Term</b>	<b>Id</b>	<b>Tf</b>	<b>Df</b>	<b>Idf</b>	<b>W</b>	<b> D </b>
Analisa	1	0.4	1	0,60206	0.240824	0.602982
Jaring	1	0.4	1	0,60206	0.240824	
komputer	1	0.4	2	0,30103	0.120412	
Kantor	1	0.4	1	0,60206	0.240824	
rektorat	1	0.4	1	0,60206	0.240824	
kampus	1	0.4	1	0,60206	0.240824	
Uin	1	0.4	2	0,30103	0.120412	
Suska	1	0.4	2	0,30103	0.120412	
Riau	1	0.4	2	0,30103	0.120412	
sistem	1	1	4	0.12493	0.120412	
sistem	2	0.4	4	0.12493	0. 240824	0.694318
Pembayaran	2	0.4	1	0,60206	0. 240824	
Uang	2	0.4	1	0,60206	0. 240824	
Bulan	2	0.4	1	0,60206	0. 240824	
sekolah	2	0.4	1	0,60206	0. 240824	
Tengah	2	0.4	1	0,60206	0. 240824	
Juru	2	0.4	1	0,60206	0. 240824	
2	2	0.4	1	0,60206	0. 240824	
pekanbaru	2	0.4	1	0,60206	0. 240824	
informasi	2	1	3	0.12493	0.124939	
optimasi	3	0.4	1	0,60206	0. 240824	0.78107
website	3	0.4	1	0,60206	0. 240824	
Uin	3	0.8	2	0,30103	0. 240824	
Suska	3	0.8	2	0,30103	0. 240824	
Riau	3	0.8	2	0,30103	0. 240824	
metode	3	0.4	1	0,60206	0. 240824	
Search	3	0.4	1	0,60206	0. 240824	
Engine	3	0.4	1	0,60206	0. 240824	
optimization	3	0.4	1	0,60206	0. 240824	
Pusat	3	0.4	1	0,60206	0. 240824	
komputer	3	0.4	2	0,30103	0.120412	1.18802
Informasi	3	1	3	0.12493	0.120412	
sistem	4	0.4	2	0,30103	0.049975	
informasi	4	0.4	1	0,60206	0.049975	
Siswa	4	1.4	1	0,60206	0.82884	
madrasah	4	0.4	1	0,60206	0. 240824	
Aliyah	4	0.4	1	0,60206	0. 240824	
Darel	4	0.8	1	0,60206	0.481648	
hikmah	4	0.8	1	0,60206	0.481648	
pondok	4	0.4	1	0,60206	0. 240824	
pesantrean	4	4	1	0,60206	0. 240824	

#### 4.2.3.2. Proses *query*

Tahapan Preproses *Query* Pengguna bisa dilihat pada Gambar 4.6



Gambar 4.6 Tahapan Preproses *Query* Pengguna

Preproses yang terjadi pada *query* secara garis besar sama dengan preproses yang terjadi pada dokumen, yaitu setelah *query* dimasukan oleh pengguna maka sistem akan:

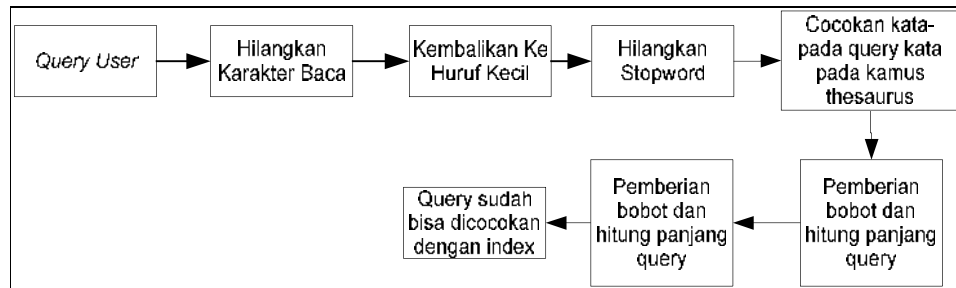
- Menghilangkan Karakter Baca dari *Query*
- Mengubah *query* Mengubah setiap huruf besar ke Huruf kecil
- Menerapkan *stopword remove* jika terdapat didalam *query*.
- Parsing dokumen dan beri bobot pada *query*, bobot dari queri diambil dari bobot yang telah ada pada *index* lalu di hitung vektor dari *query*.

Tabel 4.7 Pembobotan *Query*

Perhitungan Bobot <i>Query</i>			$idf = \log(N/df)$	$W = Tf * idf$	$ Q $
Queri	Df	Tf	Idf	W	
Sekolah	4	1	0.602059991	0.602059991	0.602059991

#### 4.2.3.3. Perluasan *query*

Jika proses *query* secara normal telah kita lihat pada Gambar 4.5 maka kita akan menjelaskan bagaimana proses perluasan *query* yang nantinya akan dilakukan pada sistem, hal ini bisa dilihat pada Gambar 4.6.



Gambar 4.7 Tahapan Preproses *Query* Pengguna dengan menerapkan perluasan pada *query*, yang diambil dari kamus thesaurus

Untuk proses perluasan hampir sama dengan proses *query* sebelumnya, akan tetapi *query* tidak langsung dilakukan pembobotan dan penghitungan vektor dari *query*, namun *query* akan dilakukan perluasan dengan mengacu dari kamus thesaurus kemudian baru diberikan bobot.

Contoh :

Ketia *query* awal dari pengguna adalah “Sekolah” Lalu dicocokkan dengan kamus Tesaurus, maka *query* akan diperluasa menjadi “Sekolah Kampus madrasah dan pesantren”

Tabel 4.8 Ilustrasi Perluasan *Query*

<i>Query</i> awal	Kamus thesaurus		<i>Query</i> Pengguna Setelah diperluas
	Kata	Perluasan	
Sekolah	Sekolah	Kampus madrasah pesantren	Sekolah kampus
			Madrasah Pesantren

Setelah selesai dilakukan Proses Pencocokan, maka *Query* dari pengguna akan dilakukan Pembobotan sesuai dengan ketentuan persamaan (2.8). dengan faktor penyesuaian 1 untuk *query* awal (Sekolah) dan 0.5 untuk *query* yang telah di perluas (kampus madrasah pesantren) untuk keterangan lebih lanjut bisa dilihat pada Tabel 4.9 .

Tabel 4.9 Ilustrasi Pembobotan *Query* Yang diberi Faktor Penyesuaian

Jenis Query	Query	Idf	W	A <sub>dj</sub>	W'
<i>Awal</i>	Sekolah	0.60205	0.60205	1	0.60205
<i>Perluasan</i>	Kampus	0.60205	0.60205	0.5	0.301029
	Madrasah	0.60205	0.60205	0.5	0.301029
	Pesantren	0.60205	0.60205	0.5	0.301029

Keterangan :

A<sub>dj</sub> = Nilai Faktor Penyesuaian yang diberikan untuk *query*

W' = Bobot *query* setelah dikalikan dengan Faktor Penyesuaian

Setelah W' didapat, maka W' akan menggantikan bobot yang lama (W) untuk digunakan didalam perhitunga *similarity* selanjutnya.

#### 4.2.3.4. Pencocokan (*similarity*)

Pencocokan atau penentuak akan kemiripan (*similarity measure*) dihitung dengan menggunakan persamaan 2.7 Maka hasil perhitungan ukuran kemiripan untuk contoh di atas dapat dilihat sebagai berikut:

Dengan keterangan :

W<sub>11</sub>=dibaca bobot kata pertama pada dokumen 1

W<sub>q1</sub>= dibaca bobot *query* pertama.

**Dok.2**

$$\begin{aligned}
 &= \frac{( ) + ( ) + ( ) + ( ) + ( ) + ( ) + ( ) + ( ) + ( ) + ( )}{[( )]} \\
 &= \frac{( . ) ( . )}{( . ) ( . ) ( . ) ( . ) ( . ) ( . ) ( . ) ( . ) ( . ) ( . )} \\
 &= \frac{.}{.}
 \end{aligned}$$

$$= \frac{0.34684971}{0.34684971}$$

$$\text{SIM}(D_1, Q) = 0.34684971$$

Dengan cara yang sama, maka diperoleh hasil *similarity*  $D_2$ ,  $D_3$  dan  $D_4$ , yang tergambar pada Tabel 4.10

Tabel 4.10 Nilai *Similarity* Setiap dokumen terhadap *query*

<b><i>SIMILARITY</i> (<math>d_i, q</math>)</b>		
<b>dok id</b>	<b><i>Similarity</i></b>	<b>Rangking</b>
dok1	0	0
dok2	<b>0.34684971</b>	1
dok3	0	0
dok4	0	0

Dari hasil perhitungan diatas, dapat dilihat bahwa dokumen 2 memiliki nilai *similarity* yang paling tinggi, Untuk dokumen 1 3 dan 4 memiliki nilai *similarity* 0. Untuk perangkian nantinya dokumen yang memiliki nilai *similarity* tertinggi akan terletak pada urutan teratas dan berurut kebawah.

Untuk perhitungan sistem setelah dilakukan perluasan adalah sebagai berikut :

Dimana *query* awal adalah sekolah dan kemudian diperluas dan menjadi Sekolah, kampus, madrasah, Pesantren. Maka perhitungan Bobotnya dengan Fakto penyesuayan 0.5 ( $W'$ ) untuk *query* hasil perluasan:

Tabel 4.11 Bobot dari *query* hasil perluasan

<b>N o</b>	<b>Term</b>	<b>D f</b>	<b>T f</b>	$Idf = \log(N/df)$		
				Idf	$W'$	$ Q $
1	Kampus	1	1	0.60205	0.301029	0.425721
2	Madrasah	1	1	0.60205	0.301029	
3	Pesantren	1	1	0.60205	0.301029	

Dengan perhitungan yang sama untuk mencari nilai *similarity* antara *query* dengan dokumen yang ada maka didapatkan nilai *similarity* yang baru :

Tabel 4.12 Nilai *Similarity* Setiap dokumen terhadap *query* setelah diperluas.

<b><i>SIMILARITY</i> (d<sub>i</sub>,q)</b>		
<b>dok id</b>	<b><i>Similarity</i></b>	<b>Rangking</b>
dok1	0.230587	3
dok2	0.34816	1
dok3	0	0
dok4	0.23407	2

### 4.3. Analisa Perancangan Sistem

Secara umum perancangan sistem terbagi kepada dua bagian utama, yakni perancangan *database* dari sistem dan yang kedua ialah perancangan antarmuka dari sistem.

#### 4.3.1. Perancangan *Database*

Didalam perancangan *database* kita akan membahas mengenai *database* yang dibutuhkan sistem.

- a. Tabel Pengguna, yang digunakan untuk menyimpan data Admin.

Tabel 4.13 Pengguna

Nama Field	Type dan Length	Primary Key	Null	Default
Username	Varchar(20)	Yes	No	-
Password	Varchar(50)		No	-

- b. Tabel *index*, yang nantinya akan digunakan untuk menyimpan hasil *index* dari dokumen yang ada.

Tabel 4.14 *index*

Nama Field	Type dan Length	Primary Key	Null	Default
Id <i>index</i>	Int(11)	Yes	No	-
Kata	Inr(25)		No	-
Id_abstrak	Varchar(20)		No	-
Jumlah	Fload		No	-
Idf	Fload		No	-
Bobot	Fload		No	-



- c. Tabel karakter, yang nantinya akan digunakan untuk menyimpan karakter-karakter baca yang ingin dihapus

Tabel 4.15 Karakter

Nama Field	Type dan Length	Primary Key	Null	Default
Id_kar	Int(10)	Yes	No	-
Nm_kar	Inr(25)		No	-

- d. Tabel *Stopword*, yang nantinya akan digunakan untuk menyimpan kata-kata *stopword* yang ingin dihapus.

Tabel 4.16 *Stopword*

Nama Field	Type dan Length	Primary Key	Null	Default
Id_index	Int(10)	Yes	No	-
Kata	varchar(25)		No	-

- d. Tabel Vektor Dokumen, yang nantinya akan digunakan untuk menyimpan Vektor dari bobot setiap dokumen yang telah dihitung.

Tabel 4.17 vektor dokumen

Nama Field	Type dan Length	Primary Key	Null	Default
Id_dok	varchar(10)	Yes	No	-
Panjang	Fload		No	-

- e. Tabel Tesaurus, yang berisikan data dari perluasan kata yang ada, yang diambil dari kamus thesaurus.

Tabel 4.18 Tesaurus

Nama Field	Type dan Length	Primary Key	Null	Default
Id_tesaurus	Int(10)	Yes	No	-
Kata	varchar(50)		No	-
Thesaurus	Varchar(500)		No	

#### 4.3.2. Perancangan Tampilan

Berikut adalah perancangan tampilan dari sistem yang dirancang agar memudahkan bagi pengguna didalam menggunakan sistem.

#### 4.3.2.1. *Form Tampilan Utama*

Pada tampilan utama ini, nantinya akan ditujukan untuk User Pengguna Umum yang ingin mendapatkan informasi

The diagram illustrates the layout of the Main Display Form. It is structured into three main horizontal sections: a top section, a middle content area, and a bottom section. The top section contains a navigation bar with a button labeled 'ABOUT' on the left and the word 'HEADER' centered. The middle section is the main content area, which is mostly empty except for a central button labeled 'MASUKAN KATA'. Below this button are two smaller buttons, 'Cari' and 'Perluasan', positioned side-by-side. The bottom section is a footer bar containing the word 'FOOTER' centered.

Gambar 4.8 Analisis *Form* tampilan utama

#### 4.3.2.2. *Form Login Admin*

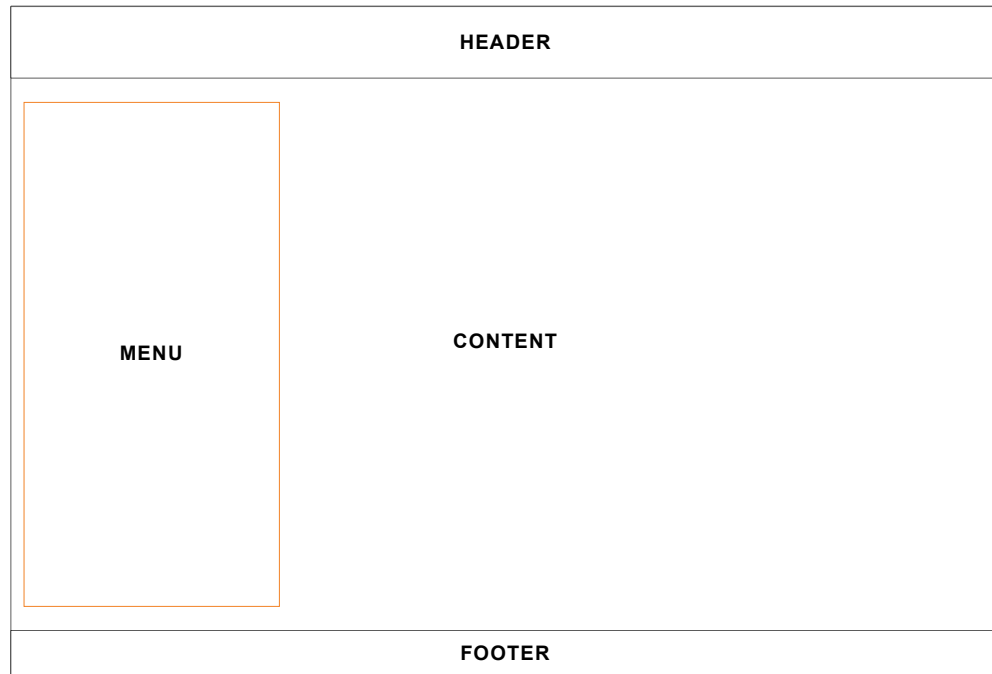
Pada sistem nantinya akan ada halaman yang akan digunakan admin untuk mengelola sistem, akan tetapi sebelum admin bisa mengelola sistem, admin harus *login* terlebih dahulu untuk bisa masuk ke halaman admin.

The diagram shows the layout of the Admin Login Form. It is a rectangular box containing two input fields. The first field is labeled 'Username' and the second is labeled 'Password'. Below these fields are two buttons: 'login' and 'reset', positioned side-by-side.

Gambar 4.9. Analisa *Form Login* Admin

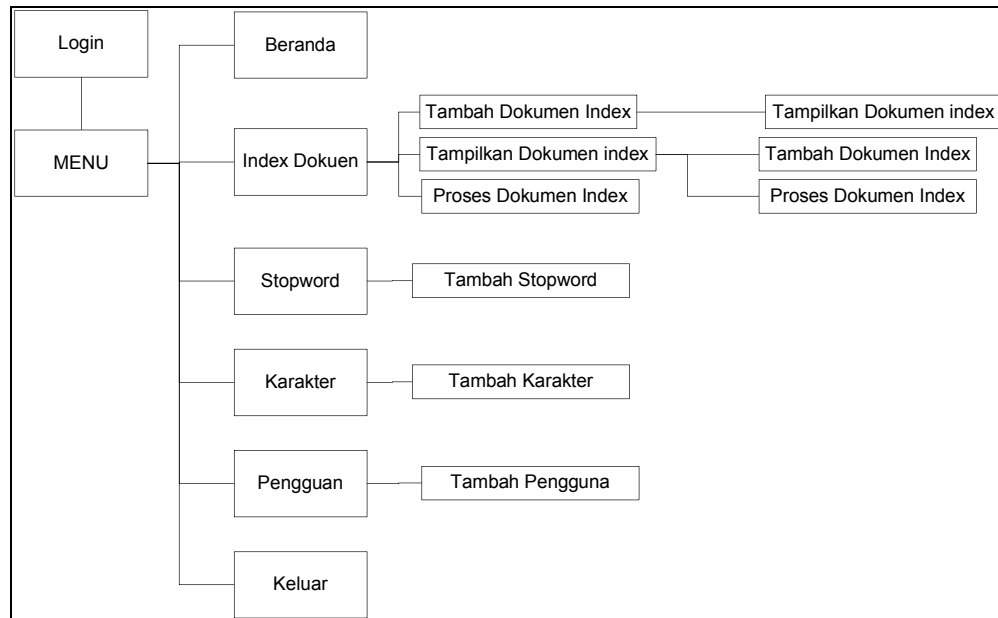
#### 4.3.2.3. *Form* Halaman Utama Admin

Setelah admin berhasil melakukan *login*, maka akan muncul halaman utama dari admin.



Gambar 4.10 Analisa *Form* Halaman Utama Admin

Didalam pengelolaan halaman admin, nantinya admin akan menemui struktur menu yang digambarkan sebagai berikut :



Gambar 4.11 Struktur Menu Admin

Untuk penjelasan dari struktur menu admin pada Gambar 4.1 bisa dilihat pada Tabel 4.18:

Tabel 4.19 Keterangan Struktur Menu Admin

NO	Nama	Keterangan ketika digunakan
1	<i>Login</i>	Sebagai keamanan dan Pintu masuk dari admin yang ingin mengelola halaman admin.
2	Menu	Menu adalah Fitur-fitur yang diberikan sistem yang bisa digunakan admin
3	Beranda	akan dikembalikan ke halaman utama.
4	Index Dokumen	Akan menampilkan hasil pengindekan sistem
5	Tambah Index Dokumen	Untuk menambahkan korpus dokumen yang ingin diindek
6	Tampilkan Dokumen Index	Menampilkan dokumen-dokumen yang telah diinputkan
7	Proses Dokumen Index	Untuk memproses dokumen menjadi indek
8	<i>Stopword</i>	untuk menampilkan data <i>Stopword</i>
9	Tambah <i>Stopword</i>	Untuk menambahkan kata <i>stopword</i>
10	Karakter	Menampilkan data karakter
11	Tambah Karakter	Untuk menambahkan karakter
12	Pengguna	Menampilkan data pengguna
13	Tambah Pengguna	Untuk menambah pengguna
14	Keluar	Untuk keluar dari sistem

## **BAB V**

### **IMPLEMENTASI DAN PENGUJIAN**

Pada bab ini akan dijelaskan implementasi dan pengujian dari sistem temu balik informasi yang telah dirancang pada bab sebelumnya. Adapun pembahasan pada tahapan implementasi meliputi batasan implementasi, lingkungan operasional dan hasil implementasi dan pada pengujian meliputi lingkungan pengujian, tahapan pengujian dan hasil pengujian.

#### **5.1 Implementasi**

Tahapan implementasi adalah kondisi dimana sistem yang telah analisa dan dirancang siap dioperasikan pada kondisi yang sebenarnya, dari tahapan implementasi ini akan diketahui tingkat keberhasilan analisa dan perancangan pada sistem yang akan dibangun.

##### **5.1.1 Batasan Implementasi**

Sistem temu balik informasi yang dibangun pada tugas akhir memiliki batasan sebagai berikut :

1. Bahasa pemrograman yang digunakan dalam pengimplementasian sistem ini yaitu Php dengan DBMS mySQL pada sistem operasi *Microsoft Windows 7 Ultimate*.
2. Perluasan *Query* yang dilakukan hanya berdasarkan kamus Tesaurus Bahasa Indonesia yang ada.
3. Koleksi Dokumen yang digunakan ialah dokumen berformat .XML yang proses *input* text dilakukan dari sistem.

### 5.1.2 Lingkungan Operasional

Adapun pengimplementasian sistem temu balik informasi ini dibagi kedalam dua komponen yaitu perangkat keras dan perangkat lunak, berikut ini adalah lingkungan operasional yang digunakan dalam pengimplementasian sistem:

#### 1. Perangkat keras

Processor : *Intel(R) Core(TM)2 Duo CPU T6670 @ 2.20GHz*  
Memori (RAM) : 1.00 GB

#### 2. Perangkat Lunak

Sistem Operasi : *Windows 7 Ultimate 32-bit (6.1, Build 7600)*  
Bahasa Pemrograman : *Php*  
DBMS : *mySQL*  
*Tools* Perancangan : *Notepad++*,

### 5.1.3 Hasil Implementasi Sistem

Adapun hasil implentasi sistem ini dibagi menjadi dua yaitu hasil implementasi perhitungan sistem dan hasil implementasi *interface* sistem.

#### 5.1.3.1. Implementasi Perhitungan Sistem

Berdasarkan perhitungan manual pada bab IV Tabel 4.6 maka pada bab implementasi ini akan dibandingkan kesesuaian perhitungan antara perhitungan secara manual dan perhitungan otomatis oleh sistem

##### 1. Dokumen Yang digunakan

Dokumen yang digunakan adalah 4 dokumen yang sama yaitu :

- a. Sistem  $d_1$  = analisa jaringan komputer pada kantor rektorat kampus UIN suska Riau
- b. Informasi  $d_2$  = sistem pembayaran uang bulanan sekolah menengah kejuruan 2 pekanbaru
- c. Informasi  $d_3$  = optimasi website uin suska riau dengan *metode search engine optimaztion* pusat computer uin suska riau

- d. Kesiswaan d<sub>4</sub> = sistem informasi kesiswaan madrasah aliyah darel hikmah pondok pesantren darel hikmah

## **2. Pembobotan dan perhitungan nilai panjang Vektor Dokumen**

Setelah dokumen berhasil disimpan maka sistem akan melakukan proses pengindekan dan pemberian bobot terhadap masing-masing kata yang mewakili dokumen.

- a. Proses Menghitung Bobot dari kata dan Panjang Vektor Dokumen 1

**Bobot Kata 1** : 0.240824

**Bobot Kata 2**: 0.240824

**Bobot Kata 3**: 0.120412

**Bobot Kata 4**: 0.240824

**Bobot Kata 5**: 0.240824

**Bobot Kata 6**: 0.240824

**Bobot Kata 7**: 0.120412

**Bobot Kata 8**: 0.120412

**Bobot Kata 9**: 0.120412

**Bobot Kata 10**: 0.124939

**Nilai Vektor Dokumen** : 0.60298171413153

- b. Proses Menghitung Bobot dari kata dan Panjang Vektor Dokumen 2

**Bobot Kata 1**: 0.0499755

**Bobot Kata 2**: 0.240824

**Bobot Kata 3**: 0.240824

**Bobot Kata 4**: 0.240824

**Bobot Kata 5**: 0.240824

**Bobot Kata 6**: 0.240824

**Bobot Kata 7**: 0.240824

**Bobot Kata 8**: 0.240824

**Bobot Kata 9**: 0.240824

**Bobot Kata 10**: 0.124939

**Nilai Vektor Dokumen** : 0.69431757584642

c. Proses Menghitung Bobot dari kata dan Panjang Vektor Dokumen 3

**Bobot Kata 1:** 0.240824

**Bobot Kata 2:** 0.240824

**Bobot Kata 3:** 0.240824

**Bobot Kata 4:** 0.240824

**Bobot Kata 5:** 0.240824

**Bobot Kata 6:** 0.240824

**Bobot Kata 7:** 0.240824

**Bobot Kata 8:** 0.240824

**Bobot Kata 9:** 0.240824

**Bobot Kata 10:** 0.240824

**Bobot Kata 11:** 0.120412

**Bobot Kata 12:** 0.124939

**Nilai Vektor Dokumen :** 0.78107028699407

d. Proses Menghitung Bobot dari kata dan Panjang Vektor Dokumen 4

**Bobot Kata 1:** 0.0499755

**Bobot Kata 2:** 0.0499755

**Bobot Kata 3:** 0.842884

**Bobot Kata 4:** 0.240824

**Bobot Kata 5:** 0.240824

**Bobot Kata 6:** 0.481648

**Bobot Kata 7:** 0.481648

**Bobot Kata 8:** 0.240824

**Bobot Kata 9:** 0.240824

**Nilai Vektor Dokumen :** 1.1880248004013

Setelah Proses diatas selesai, maka indek pada sistem bisa dinyatakan telah selesai dan bisa digunakan, selanjutnya ialah menguji hasil dari perhitungan Proses dari *input query* pengguna sehingga dadapatlah nilai *smilarity* antara *query* dengan dokumen.



### 3. Pembobotan *query* dan menghitung panjang vektor *query* Tanpa Perluasan

Adapun *query* yang diinputkan yaitu "Sekolah" maka dengan menggunakan persamaan didapatkan nilai *query* sebagai berikut :

**Bobot *Query* :** 0.602059

**Panjang Vektor *Query* :** 0.602059

### 4. Perhitungan *Similarity* dengan *Query* awal

Tahapan terakhir adalah melakukan perhitungan *Similarity* maka didapatkan perhitungan sebagai berikut :

**Id Dok :** dokumen2

**Kata Yang Cocok dengan *Query* :** Sekolah

**Bobot Kata pada Dokumen :** 0.240824

**Nilai *Similarity* :** 0.3468497

Sehingga yang dikembalikan kepada pengguna hanyalah **dokumen 2**

### 5. Pembobotan *query* dan menghitung panjang vektor *query* Dengan Perluasan *Query*

Dengan dilakukan perluasan maka *query* menjadi mendapatkan tambahan, yaitu pesantren madrasah dan kampus, jika *Query* ini diinputkan langsung oleh pengguna dan dilakukan perhitungan seperti biasa maka bobot yang didapat sebagai berikut :

**Bobot *Query* 1 :** 0.602059

**Bobot *Query* 2 :** 0.602059

**Bobot *Query* 3 :** 0.602059

**Nilai Vektor *Query* :** 1.0427984941845

Akan tetapi, dengan proses perluasan *query* yang menerapkan konsep pemberian nilai faktor penyesuaian sebesar 0.5 untuk *query* hasil perluasan, maka bobot yang didapat sebagai berikut :

**Bobot *Query* 1:** 0.30102999566398

**Bobot *Query* 2:** 0.30102999566398

**Bobot *Query* 3 :** 0.30102999566398

**Nilai Vektor *Query* :** 0.521399

#### 4. Perhitungan *Similarity* dengan *Query* yang telah diperluas

Maka hasil perhitungan similarity dengan *query* yang diperluasa sebagai berikut :

**Id Dok :** dokumen1

**Kata yang Cocok dengan *Query* :** kampus

**Bobot kata :** 0.240824

**Menghitung Nilai *Similarity* dokumen1 :** 0.230586

**Id Dok :** dokumen4

**Kata yang Cocok dengan *Query* :** madrasah

**Bobot kata :** 0.240824

**Id Dok :** dokumen4

**Kata yang Cocok dengan *Query* :** pesantren

**Bobot kata :** 0.240824

**Menghitung Nilai *Similarity* dokumen4 :** 0.234069

Sehingga hasil pencarian yang dikembalikan kepada pengguna menjadi bertambah banyak, hal ini bisa dilihat pada tabel 5.1

Table 5.1 Hasil pencarian setelah diperluas

<b>Rangking</b>	<b>Id Dokumen</b>	<b><i>Nilai Similarity</i></b>
1	Dokumen 2	0.3468497
2	Dokumen 4	0.234069
3	Dokumen 1	0.230586

#### 5.1.3.2. Implementasi *Interface* Sistem

Adapun hasil implementasi *interface* sistem temu balik informasi ini mengacu pada rancangan *interface* pada bab analisa dan perancangan.

### a. Implementasi Tampilan Menu Utama Sistem

Gambar 5.1 adalah menu utama dari sistem yang dibangun, berawal dari menu utama inilah Admin dapat memilih menu dan mengoperasikan sistem



Gambar 5.1 Menu Utama Admin

### b. Implementasi *Form* Koleksi Dokumen

Gambar 5.2 adalah *form* koleksi untuk menambahkan dokumen baru pada sistem temu balik informasi.

Kode Dokumen	<input type="text" value="dokumen5"/>
Judul	<input type="text"/>
Pemilik	<input type="text"/>
Tanggal	<input type="text"/>
Abstrak	<div></div>
<div><input type="button" value="Simpan"/> <input type="button" value="Reset"/></div>	

Gambar 5.2 *Form Koleksi Dokumen Sistem*

### c. Hasil dari koleksi dokumen

Hasil dari *input* koleksi dokumen dari Gambar 5.2 menghasilkan Gambar 5.3

C:\xampp\htdocs\Query-expansion\admin\corpus.xml - Notepad++

File Edit Switch View Encoding Language Settings Macro Run TestFX Plugins Window ?

proses.php index.html home.php baranda.php cok1.xml

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <doc>
3      <docno>dokumen1</docno>
4      <title>sistem</title>
5      <author>khairi</author>
6      <date>jun 1 2011</date>
7      <text>analisa jaringan komputer pada: kantor rektorat kampus UTN suska Riau.</text>
8  </doc>

```

Gambar 5.3 Hasil *Input* Koleksi Dokumen

#### d. Implementasi Menampilkan Koleksi Dokumen

Stelah dokumen berhasil dibentuk menjadi korpus berformat .XML maka admin bisa melihat daftar dokumen yang telah di inputkan.

[Tambah Data Tugas Akhir](#)  
[Preprocessing Data Tugas Akhir](#)

ID	Judul	Pemilik	Tanggal	Abstrak
dokumen1	sistem	khairi	jun 1 2011	analisa jaringan komputer pada: kantor rektorat kampus UIN suska Riau.
dokumen2	informasi	indah	jun 1 2013	sistem Pembayaran uang bulanan sekolah menengah kejuruan 2 Pekanbaru
dokumen3	informasi	Reni	mar 2 2011	optimasi website uin suska riau dengan metode search engine optimization pusat komputer uin suska riau
dokumen4	kesiswaan	siswa	23-oktobe 2012	sistem informasi kesiswaan madrasah aliyah darel hikmah pondok pesantren darel hikmah

Gambar 5.4 Menampilkan Koleksi Dokumen Sistem

#### e. Implementasi Menampilkan Hasil Pengindekan

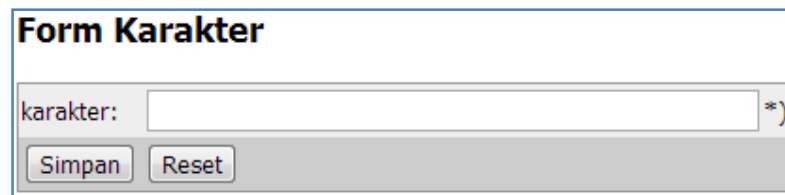
Setelah proses penyimpanan koleksi dokumen selesai, maka admin bisa melakukan pengindekan dokumen, dan hasilnya bisa dilihat pada tabel hasil pengindekan.

<a href="#">Tambah Dokumen Index</a> <a href="#">Tampilkan Dokumen Index</a> <a href="#">Proses Dokumen Index</a>						
No	Kata	Id Abstrak	Jumlah	Idf	Bobot	
1	analisa	dokumen1	0.4	0.60206	0.240824	
2	jaringan	dokumen1	0.4	0.60206	0.240824	
3	komputer	dokumen1	0.4	0.30103	0.120412	
4	kantor	dokumen1	0.4	0.60206	0.240824	
5	rektorat	dokumen1	0.4	0.60206	0.240824	
6	kampus	dokumen1	0.4	0.60206	0.240824	
7	uin	dokumen1	0.4	0.30103	0.120412	
8	suska	dokumen1	0.4	0.30103	0.120412	
9	riau	dokumen1	0.4	0.30103	0.120412	
10	sistem	dokumen1	1	0.124939	0.124939	
11	sistem	dokumen2	0.4	0.124939	0.0499755	
12	pembayaran	dokumen2	0.4	0.60206	0.240824	
13	uang	dokumen2	0.4	0.60206	0.240824	
14	bulanan	dokumen2	0.4	0.60206	0.240824	
15	sekolah	dokumen2	0.4	0.60206	0.240824	
16	menengah	dokumen2	0.4	0.60206	0.240824	
17	kejuruan	dokumen2	0.4	0.60206	0.240824	
18	2	dokumen2	0.4	0.60206	0.240824	
19	pekanbaru	dokumen2	0.4	0.60206	0.240824	
20	informasi	dokumen2	1	0.124939	0.124939	
21	optimasi	dokumen3	0.4	0.60206	0.240824	
22	website	dokumen3	0.4	0.60206	0.240824	

Gambar 5.5 Menampilkan Hasil Pengindekan

**f. Implementasi *Form input Stop Character***

Gambar 5.6 adalah *form* untuk menambahkan daftar karakter yang akan digunakan sebagai filter karakter pada koleksi dokumen.



The image shows a web form titled "Form Karakter". It contains a label "karakter:" followed by a text input field. To the right of the input field is a small icon of a smiley face with asterisks "\*)". Below the input field are two buttons: "Simpan" (Save) and "Reset".

Gambar 5.6 *Form input Stop* karakter

**g. Implementasi menampilkan Karatet**

Setelah data karakter berhasil ditambahkan maka admin bisa melihat lis dari karakter yang telah ada pada Gambar 5.7

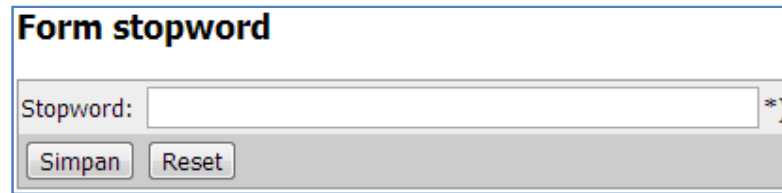
Tambah Karakter

No ↕	Karakter↕	Hapus↕	Edit↕
1	'	<a href="#">Edit</a>	<a href="#">Hapus</a>
2	-	<a href="#">Edit</a>	<a href="#">Hapus</a>
4	)	<a href="#">Edit</a>	<a href="#">Hapus</a>
5	\	<a href="#">Edit</a>	<a href="#">Hapus</a>
6	/	<a href="#">Edit</a>	<a href="#">Hapus</a>
7	=	<a href="#">Edit</a>	<a href="#">Hapus</a>
8	,	<a href="#">Edit</a>	<a href="#">Hapus</a>
9	.	<a href="#">Edit</a>	<a href="#">Hapus</a>
10	:	<a href="#">Edit</a>	<a href="#">Hapus</a>
11	;	<a href="#">Edit</a>	<a href="#">Hapus</a>
12	!	<a href="#">Edit</a>	<a href="#">Hapus</a>
13	?	<a href="#">Edit</a>	<a href="#">Hapus</a>
14	©	<a href="#">Edit</a>	<a href="#">Hapus</a>
15	"	<a href="#">Edit</a>	<a href="#">Hapus</a>

Gambar 5.7 Menampilkan Lis Karakter yang telah diinputkan

**h. Implementasi *Form input Stop Word***

Gambar 5.8 adalah *form* untuk menambahkan daftar kata yang akan digunakan sebagai filter kata pada koleksi dokumen.



**Form stopwords**

Stopword:  \*

Gambar 5.8 *Form input Stop Word*

**i. Implementasi menampilkan Daftar Stopword**

Setelah data Stopword berhasil ditambahkan maka admin bisa melihat lis dari Stopword yang telah ada pada Gambar 5.9

Tambah stopwords

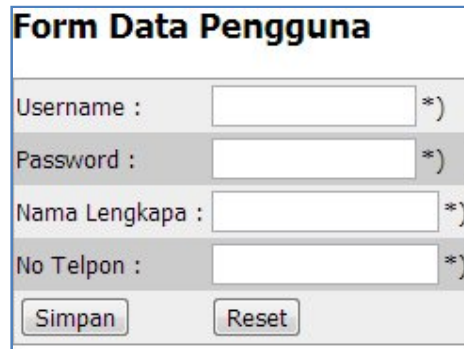
No	Kata	Hapus	Edit
1	yang	<a href="#">Edit</a>	<a href="#">Hapus</a>
2	di	<a href="#">Edit</a>	<a href="#">Hapus</a>
3	dan	<a href="#">Edit</a>	<a href="#">Hapus</a>
4	itu	<a href="#">Edit</a>	<a href="#">Hapus</a>
5	dengan	<a href="#">Edit</a>	<a href="#">Hapus</a>
6	untuk	<a href="#">Edit</a>	<a href="#">Hapus</a>
7	tidak	<a href="#">Edit</a>	<a href="#">Hapus</a>
8	ini	<a href="#">Edit</a>	<a href="#">Hapus</a>
9	dari	<a href="#">Edit</a>	<a href="#">Hapus</a>
10	dalam	<a href="#">Edit</a>	<a href="#">Hapus</a>
11	akan	<a href="#">Edit</a>	<a href="#">Hapus</a>
12	pada	<a href="#">Edit</a>	<a href="#">Hapus</a>
13	juga	<a href="#">Edit</a>	<a href="#">Hapus</a>
14	saya	<a href="#">Edit</a>	<a href="#">Hapus</a>
15	ke	<a href="#">Edit</a>	<a href="#">Hapus</a>
16	karena	<a href="#">Edit</a>	<a href="#">Hapus</a>
17	tersebut	<a href="#">Edit</a>	<a href="#">Hapus</a>
18	bisa	<a href="#">Edit</a>	<a href="#">Hapus</a>
19	ada	<a href="#">Edit</a>	<a href="#">Hapus</a>
20	mereka	<a href="#">Edit</a>	<a href="#">Hapus</a>

1 2 3 > >>

Gambar 5.9 Menampilkan Lis Karakter yang telah diinputkan

**j. Implementasi *Form input* Pengguna**

Gambar 5.10 adalah *form* untuk menambahkan data pengguna/admin yang bisa *Login* kedalam sistem.



**Form Data Pengguna**

Username :  \*)

Password :  \*)

Nama Lengkap :  \*)

No Telpn :  \*)

Gambar 5.10 *Form inout* Pengguna

**k. Implementasi menampilkan Daftar Pengguna yang bisa *Login***

Setelah Data pengguna berhasil disimpan maka admin bisa melihat lis pengguna yang telah ada

[Tambah Pengguna](#)

NO	USERNAME	NAMA LENGKAP	NO TELPON	Edit	HAPUS
1	admin	mimin	0831312333	<a href="#">Edit</a>	<a href="#">Hapus</a>
2	Hairi	Khairi Lestari	085264063331	<a href="#">Edit</a>	<a href="#">Hapus</a>

Gambar 5.11 Menampilkan Lis pengguna yang telah diinputkan

**l. Implementasi *Form* Pencarian STBI**

Gambar 5.12 adalah *form* pencarian STBI yang dapat diinputkan *Query* oleh pengunjung, berdasarkan *Query* yang diinputkan tersebut maka sistem akan menemubalikkan dokumen relevan yang teranking.



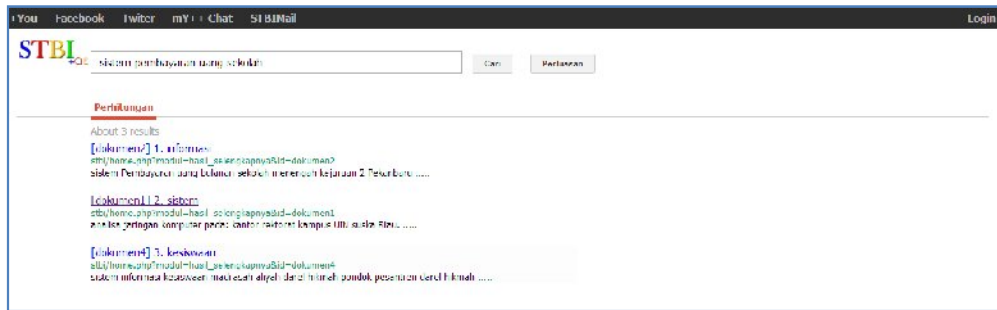
STBI +QE

Gambar 5.12 *Form* Pencarian STBI



### m. Implementasi Hasil Pencarian

Setelah *query* pengguna diproses oleh sistem maka hasil pencarian akan ditampilkan sebagai berikut :



Gambar 5.13 Hasil Pencarian

### n. Implementasi hasil pencarian selengkapnya

Setelah hasil pencarian didapat, maka isi dari dokumen yang ada bisa dilihat dengan lengkap, dengan mengklik dokumen mana yang ingin dilihat.



Gambar 5.14 Hasil Pencarian Selengkapnya

## 5.2 Pengujian

Untuk melakukan pengujian sistem dalam menghitung tingkat performansi sistem temu balik informasi menggunakan pengujian *recall* dan *precision*. Berdasarkan hasil dari pengujian *recall* dan *precision* tersebut, tingkat kelayakan dari sistem yang dibangun dapat diketahui.

### 5.2.1 Ruang Lingkup Pengujian

Adapun pengujian sistem temu balik informasi ini dibagi kedalam dua komponen yaitu perangkat keras dan perangkat lunak, berikut ini adalah lingkungan operasional yang digunakan dalam pengujian sistem:

#### 1 Perangkat keras

Processor : *Intel(R) Core(TM)2 Duo CPU T6670 @ 2.20GHz*  
Memori (RAM) : 1.00 GB

#### 2 Perangkat Lunak

Sistem Operasi : *Windows 7 Ultimate 32-bit (6.1, Build 7600)*  
Bahasa Pemrograman : *Php*  
DBMS : *mySQL*  
Tools Perancangan : *Notepad++*

### 5.2.2 Rencana Pengujian

Adapun rencana pengujian yang akan dilakukan sebagai berikut :

1. Jumlah *Query* yang akan dilakukan pengujian sebanyak 3 *Query* yang berbeda dan kemudian 3 *query* perluasan yang dilakukan oleh sistem.
2. Jumlah dokumen yang akan dilakukan pengujian sebanyak 120 dokumen.
3. Pengujian akan dilakukan dua kali, yang pertama pengujian tanpa *query expansion* dan yang kedua pengujian dengan *query expansion*
4. Untuk pengambilan hasil *retrieval* yang akan dihitung yaitu maksimal 20 *Retrieval*.
5. Dari pengujian *Query* pada dokumen maka akan didapatkan nilai *recall* dan *precision*, yang nantinya sebagai tolok ukur dari kualitas *Retrieval*.

### 5.2.3 Hasil Pengujian

Pengujian Kerelevanan Hasil Pencarian *Query Expansion* Pada Sistem Temu Kembali Informasi Dengan Model Ruang Vektor diambil berdasarkan pendapat tiga orang ahli, yang menyatakan sebuah hasil pencarian dikatakan relevan atau tidak, jika telah ada dua orang ahli menyatakan sebuah hasil

pencarian relevan maka hasilnya dinyatakan relevan begitu juga sebaliknya untuk menyatakan sebuah hasil pencarian dinyatakan tidak relevan.

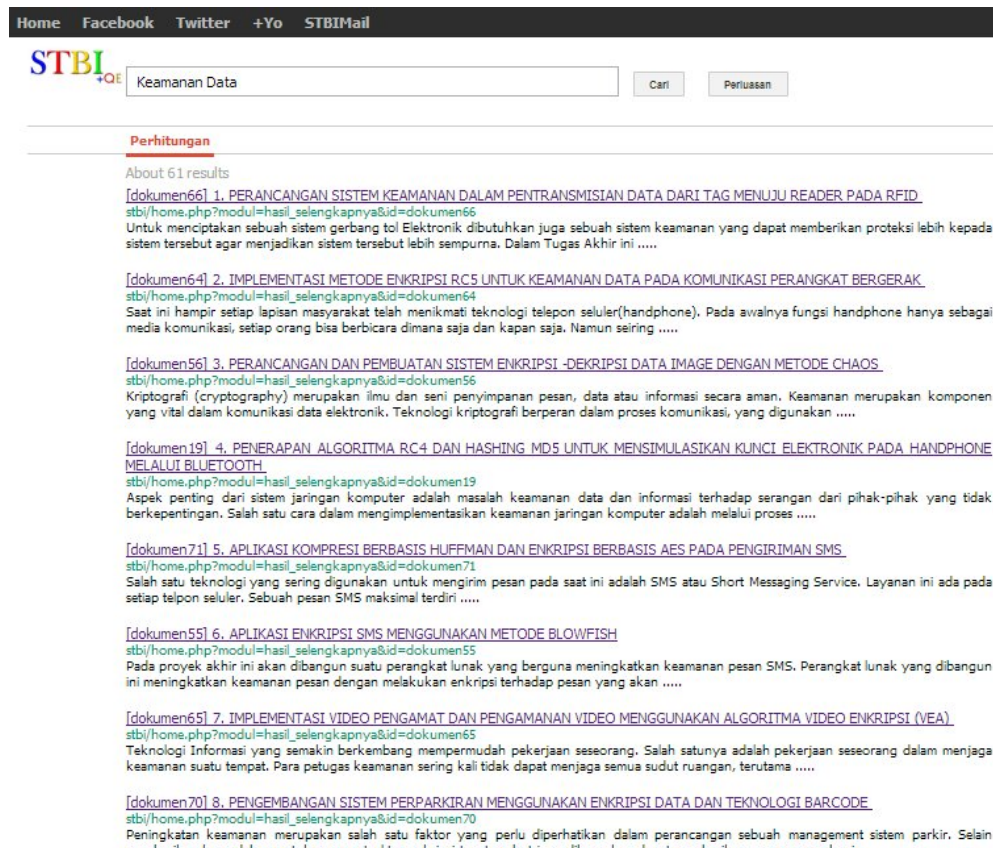
Untuk hasil pengujian selengkapnya berdasarkan pendapat para ahli bisa dilihat pada lampiran A. Berikut ini adalah *query* yang akan digunakan sebagai inputan Pengujian :

Tabel 5.2 Daftar *query* Pengujian

<i>NO</i>	<i>Query</i> Tanpa Perluasan	<i>Query</i> dengan Perluasan
1	Keamanan Data	Keamanan Data Keselamatan Informasi
2	Komunikasi Mobile	Komunikasi mobile koneksi
3	Pendukung keputusan	Pendukung keputusan penunjang kesimpulan pertimbangan

#### 5.2.3.1. Pengujian *Query* 1

Setelah *query* diinputkan, sistem akan menghitung relevansi dengan koleksi dokumen yang telah ada. Berdasarkan perhitungan rumus *recall* dan *precision* pada rumus 2.7 dan rumus 2.8 Pada pengujian dengan *input query* (Keamanan Data) Hasil pencari dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.15 dan Tabel 5.3



Gambar 5.15 Pengujian Hasil *Query*

Tabel 5.3 Perhitungan *recall* dan *Presicion (Q1)* Tanpa Perluasan

No	Id Dokumen	Relevan?	<i>Precision (P)</i>	<i>Recall (R)</i>
1	66	Relevan	1/1= 1	1/19= 0.052632
2	64	Relevan	2/2= 1	2/19= 0.105263
3	56	Relevan	3/3= 1	3/19= 0.157895
4	19	Relevan	4/4= 1	4/19= 0.210526
5	71	Relevan	5/5= 1	5/19= 0.263158
6	55	Relevan	6/6= 1	6/19= 0.315789
7	65	Relevan	7/7= 1	7/19= 0.368421
8	70	Relevan	8/8= 1	8/19= 0.421053
9	98	Relevan	9/9= 1	9/19= 0.473684
10	68	Relevan	10/10= 1	10/19= 0.526316
11	72	Relevan	11/11= 1	11/19= 0.578947
12	57	Relevan	12/12= 1	12/19= 0.631579

Tabel 5.3 Lanjutan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
13	73	Relevan	13/13= 1	13/19= 0.684211
14	52	Relevan	14/14= 1	14/19= 0.736842
15	92	Relevan	15/15= 1	15/19= 0.789474
16	109	Relevan	16/16= 1	16/19= 0.842105
17	69	Relevan	17/17= 1	17/19= 0.894737
18	61	Relevan	18/18= 1	18/19= 0.947368
19	63	Relevan	19/19= 1	19/19= 1
20	107	Tidak Relevan	19/20= 0.95	19/19= 1

### 5.2.3.2. Pengujian *Query* 1 dengan Perluasan

Pada pengujian dengan *input query* Q<sub>1</sub> (Keamanan Data) dengan Perluasan, maka *query* bertambah menjadi (Keamanan, keselamatan) dan (Data , informasi). Hasil pencarian dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.16 dan Tabel 5.4

Home Facebook Twitter +Yo STBIMail

**STBI** Query

Keamanan Data

---

**Perhitungan**

About 84 results

[dokumen107] 1. PERANCANGAN SISTEM INFORMASI PENGOLAHAN DATA PELANGGAN PADA PERUSAHAAN DAERAH AIR MINUM (PDAM) KOTA PEKANBARU  
[stbi/home.php?modul=hasil\\_selengkapnya&id=dokumen107](#)  
 Permasalahan yang sering muncul dalam pengolahan data pelanggan PDAM yaitu masih kurangnya usaha untuk memberikan berbagai informasi kepada pelanggannya. Hal ini terjadi karena kompleksitas dalam pengolahan data pelanggan meliputi jumlah .....

[dokumen30] 2. LOCATION BASED ALERT PADA MOBILE PHONE BERBASIS ANDROID  
[stbi/home.php?modul=hasil\\_selengkapnya&id=dokumen30](#)  
 Agenda di mobile phone menjadi salah satu aplikasi yang sering digunakan oleh para pengguna mobile phone untuk menyimpan segala kegiatan yang akan dilakukan dalam kurun waktu tertentu. Namun, seringkali Agenda .....

[dokumen34] 3. PEMBUATAN APLIKASI MOBILE CLIENT TRAFFIC REPORT SYSTEM BERBASIS SISTEM OPERASI ANDROID  
[stbi/home.php?modul=hasil\\_selengkapnya&id=dokumen34](#)  
 Tingkat kepadatan arus lalu-lintas di daerah perkotaan yang tidak pasti dan kerap berubah secara drastis mengakibatkan kekhawatiran dari pengguna kendaraan bermotor di jalan atas kemungkinan munculnya kemacetan arus lalu-lintas. Bagi .....

[dokumen19] 4. PENERAPAN ALGORITMA RC4 DAN HASHING MD5 UNTUK MENSIMULASIKAN KUNCI ELEKTRONIK PADA HANDPHONE MELALUI BLUETOOTH  
[stbi/home.php?modul=hasil\\_selengkapnya&id=dokumen19](#)  
 Aspek penting dari sistem jaringan komputer adalah masalah keamanan data dan informasi terhadap serangan dari pihak-pihak yang tidak berkepentingan. Salah satu cara dalam mengimplementasikan keamanan jaringan komputer adalah melalui proses .....

[dokumen111] 5. SISTEM TEMU BALIK INFORMASI (STBI) LAPORAN KERJA PRAKTEK DAN TUGAS AKHIR MENGGUNAKAN MODEL RUANG VEKTOR (STUDI KASUS : TEKNIK INFORMATIKA)  
[stbi/home.php?modul=hasil\\_selengkapnya&id=dokumen111](#)  
 Sistem temu balik informasi merupakan sebuah sistem pencarian yang menemukbalikan dokumen relevan sesuai dengan kata kunci yang diinputkan pengguna, sistem ini bekerja pada informasi yang tidak terstruktur dengan latar pengetahuan .....

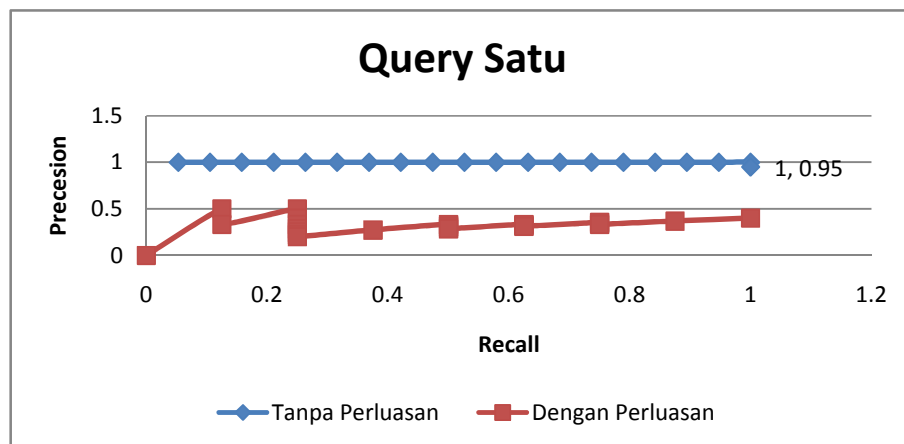
[dokumen114] 6. RANCANG BANGUN APLIKASI MOBILE KULINER (eMKUL) MENGGUNAKAN LOCATION BASED SERVICE (LBS) BERBASIS

Gambar 5.16 Pengujian Pengujian Q<sub>1</sub> Dengan Perluasan

Tabel 5.4 Perhitungan *recall* dan *Precision* (Q1) Dengan Perluasan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
1	107	Tidak Relevan	0/1= 0	0/8= 0
2	30	Relevan	1/2= 0.5	1/8= 0.125
3	34	Tidak Relevan	1/3= 0.333333	1/8= 0.125
4	19	Relevan	2/4= 0.5	2/8= 0.25
5	111	Tidak Relevan	2/5= 0.4	2/8= 0.25
6	114	Tidak Relevan	2/6= 0.333333	2/8= 0.25
7	17	Tidak Relevan	2/7= 0.285714	2/8= 0.25
8	78	Tidak Relevan	2/8= 0.25	2/8= 0.25
9	108	Tidak Relevan	2/9= 0.222222	2/8= 0.25
10	2	Tidak Relevan	2/10= 0.2	2/8= 0.25
11	66	Relevan	3/11= 0.272727	3/8= 0.375
12	84	Relevan	4/12= 0.333333	4/8= 0.5
13	12	Tidak Relevan	4/13= 0.307692	4/8= 0.5
14	31	Tidak Relevan	4/14= 0.285714	4/8= 0.5
15	56	Relevan	5/15= 0.333333	5/8= 0.625
16	52	Tidak Relevan	5/16= 0.3125	5/8= 0.625
17	64	Relevan	6/17= 0.352941	6/8= 0.75
18	90	Tidak Relevan	6/18= 0.333333	6/8= 0.75
19	70	Relevan	7/19= 0.368421	7/8= 0.875
20	65	Relevan	8/20= 0.4	8/8= 1

Data interpolasi *recall* dan *precision* pada Tabel 5.3 dan Tabel 5.4 dapat digambarkan secara grafik Pada Gambar 5.17

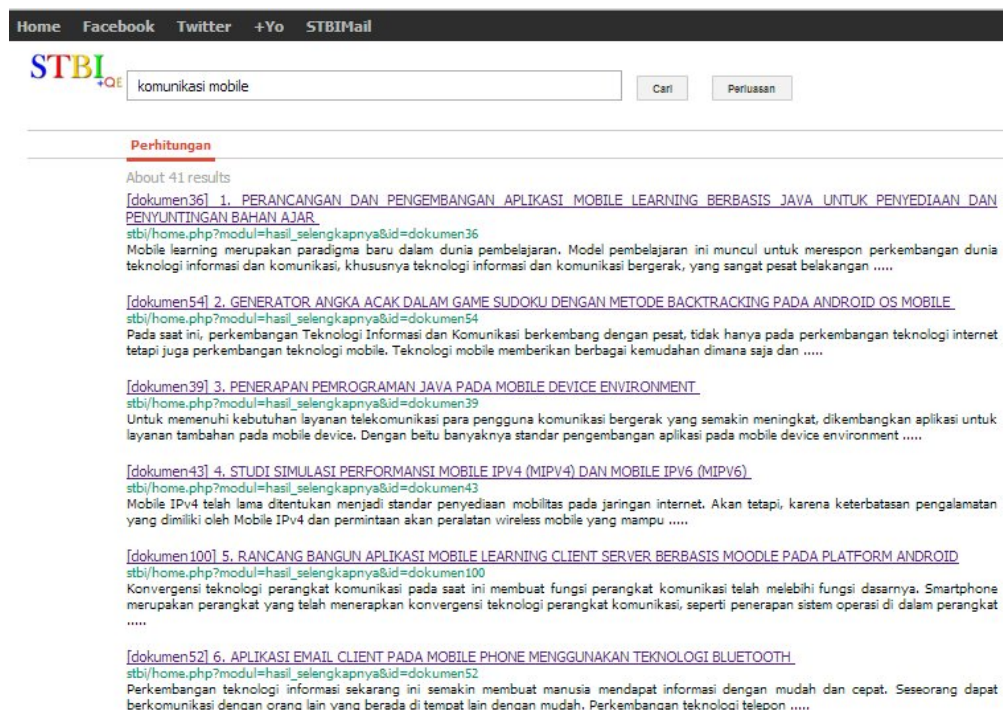


Gambar 5.17 Interpolasi *Recall* dan *Precision*  $Q_1$  Tanpa perluasan

Dimana pada Gambar 5.17 diatas terlihat bahwa nilai *Precesion* pada pengujian tanpa perluasan, lebih stabil jika dibanding dengan pengujian dengan perluasan. Contohnya pada titik pertama pada pencarian tanpa perluasan nilai *precasion* sama dengan 1 sedangkan hasil pencarian dengan perluasan bernilai 0. begitu juga dengan titik kelima, dimana nilai *precasion* tanpa perluasan samadengan 1 sedangkan dengan perluasan bernilai 0,4. Untuk perbandingan selengkapnya bisa langsung dilihat pada Tabel 5.3, Tabel 5.4 dan juga pada Gambar 5.17.

### 5.2.3.3. Pengujian *Query* 2

Pengujian dengan *input query*  $Q_2$  (komunikasi mobile). Hasil pengujian dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.18 dan Tabel 5.5



Gambar 5.18 Hasil Pengujian  $Q_2$

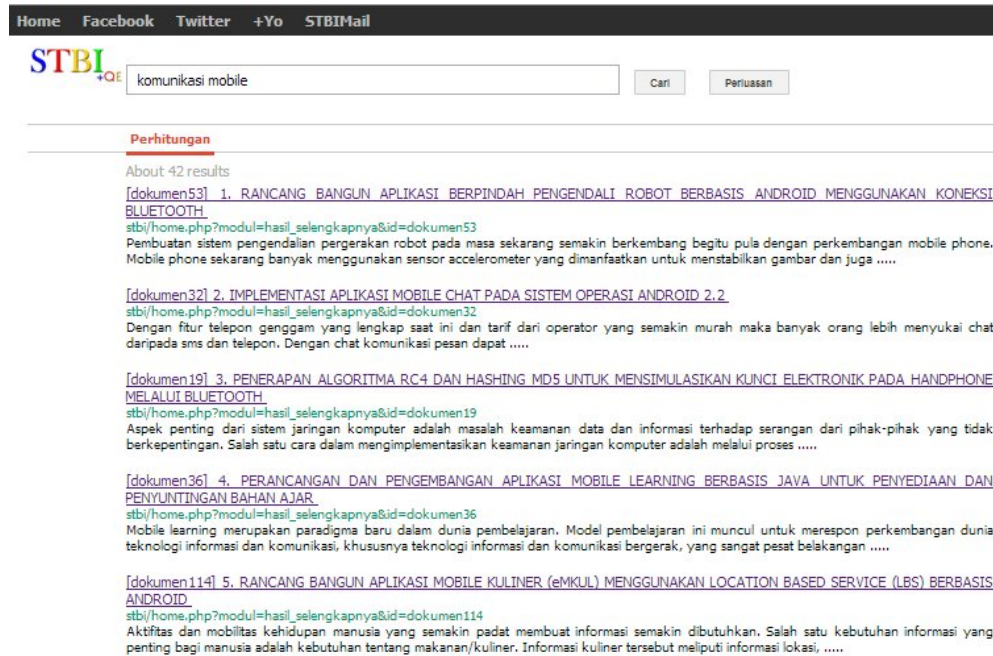
Tabel 5.5 Perhitungan *recall* dan *Precision* ( $Q_2$ ) Tanpa Perluasan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
1	36	Relevan	1/1= 1	1/19= 0.052632
2	54	Relevan	2/2= 1	2/19= 0.105263
3	39	Relevan	3/3= 1	3/19= 0.157895
4	43	Relevan	3/4= 1	3/19= 0.210526
5	100	Relevan	5/5= 1	5/19= 0.263158
6	52	Relevan	6/6= 1	6/19= 0.315789
7	41	Relevan	7/7= 1	7/19= 0.368421
8	49	Relevan	8/8= 1	8/19= 0.421053
9	33	Relevan	9/9= 1	9/19= 0.473684
10	44	Relevan	10/10= 1	10/19= 0.526316
11	48	Relevan	11/11= 1	11/19= 0.578947
12	37	Relevan	12/12= 1	12/19= 0.631579
13	38	Relevan	13/13= 1	13/19= 0.684211
14	31	Relevan	14/14= 1	14/19= 0.736842
15	45	Relevan	15/15= 1	15/19= 0.789474
16	32	Relevan	16/16= 1	16/19= 0.842105
17	47	Relevan	17/17= 1	17/19= 0.894737
18	30	Relevan	18/18= 1	18/19= 0.947368
19	64	Tidak	18/19= 0.947368	18/19= 0.947368
20	40	Ya	19/20= 0.95	19/19= 1

#### 5.2.3.4. Pengujian *Query* 2 Dengan Perluasan

Pada pengujian dengan *input query*  $Q_2$  (komunikasi *mobile*) dengan Perluasa, maka *query* bertambah menjadi (komunikasi, koneksi, dan *mobile*). Hasil pengujian dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.19 dan Tabel 5.6





Gambar 5.19 Hasil Pengujian  $Q_2$  Dengan Perluasan

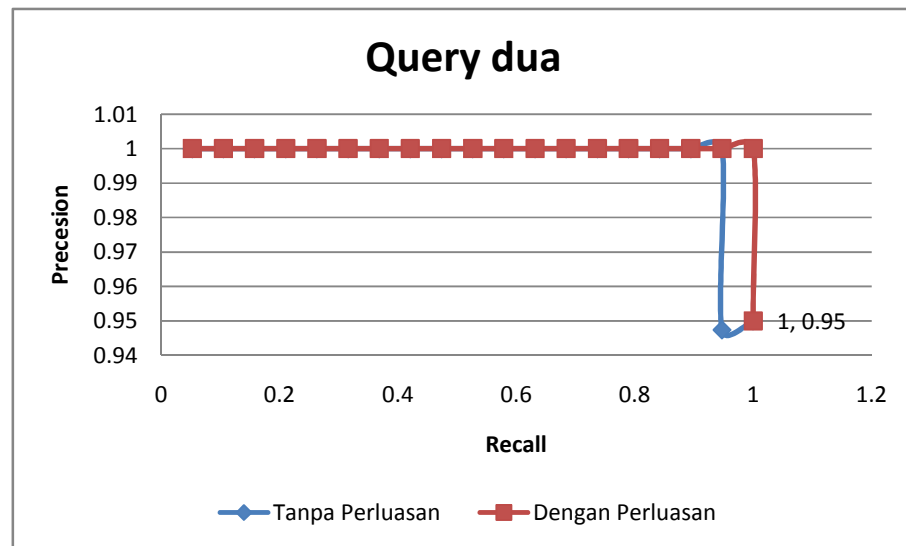
Tabel 5.6 Perhitungan *recall* dan *Precision* ( $Q_2$ ) Dengan Perluasan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
1	53	Relevan	1/1= 1	1/19= 0.05263
2	32	Relevan	2/2= 1	2/19= 0.10526
3	19	Relevan	3/3= 1	3/19= 0.15789
4	36	Relevan	4/4= 1	4/19= 0.21053
5	114	Relevan	5/5= 1	5/19= 0.26316
6	54	Relevan	6/6= 1	6/19= 0.31579
7	39	Relevan	7/7= 1	7/19= 0.36842
8	43	Relevan	8/8= 1	8/19= 0.42105
9	100	Relevan	9/9= 1	9/19= 0.47368
10	52	Relevan	10/10= 1	10/19= 0.52632
11	41	Relevan	11/11= 1	11/19= 0.57895
12	49	Relevan	12/12= 1	12/19= 0.63158
13	33	Relevan	13/13= 1	13/19= 0.68421
14	44	Relevan	14/14= 1	14/19= 0.73684
15	48	Relevan	15/15= 1	15/19= 0.78947
16	37	Relevan	16/16= 1	16/19= 0.84211

Tabel 5.6 Lanjutan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
17	38	Relevan	17/17= 1	17/19= 0.89474
18	31	Relevan	18/18= 1	18/19= 0.94737
19	45	Relevan	19/19= 1	19/19= 1
20	47	Tidak Relevan	19/20= 0.95	19/19= 1

Data interpolasi *recall* dan *precision* pada Tabel 5.5 dan Tabel 5.6 dapat digambarkan secara grafik Garis

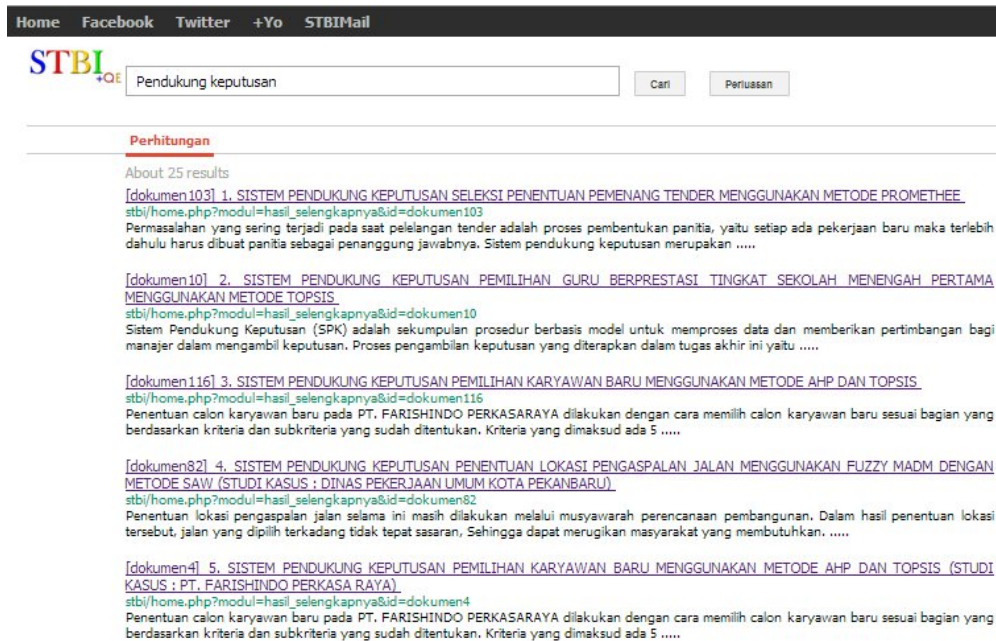


Gambar 5.20 Interpolasi *Recall* dan *Precision* Q<sub>2</sub> Tanpa perluasan

Dimana pada Gambar 5.20 diatas terlihat bahwa perbandingan nilai *Precision* pada pengujian terlihat hamper sama, pada titik sembilanbelas nilai *precision* pengujian dengan perluasan lebih tinggi yakni 1 jika dibandingkan dengan tanpa perluasan yang bernilai 0,94 sedangkan pada titik duapuluh nilai *precision* keduanya sama yakni 0.95 dan Untuk perbandingan selengkapnya bisa langsung dilihat pada Tabel 5.5, Tabel 5.6 dan juga pada Gambar 5.20.

### 5.2.3.5. Pengujian Query 3

Pengujian dengan *input query* Q<sub>3</sub> (Pendukung keputusan), Hasil Pengujian dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.21 dan Tabel 5.7



Gambar 5.21 Hasil Pengujian Q<sub>3</sub>

Tabel 5.7 Perhitungan *recall* dan *Presicion* (Q<sub>3</sub>) Tanpa Perluasan

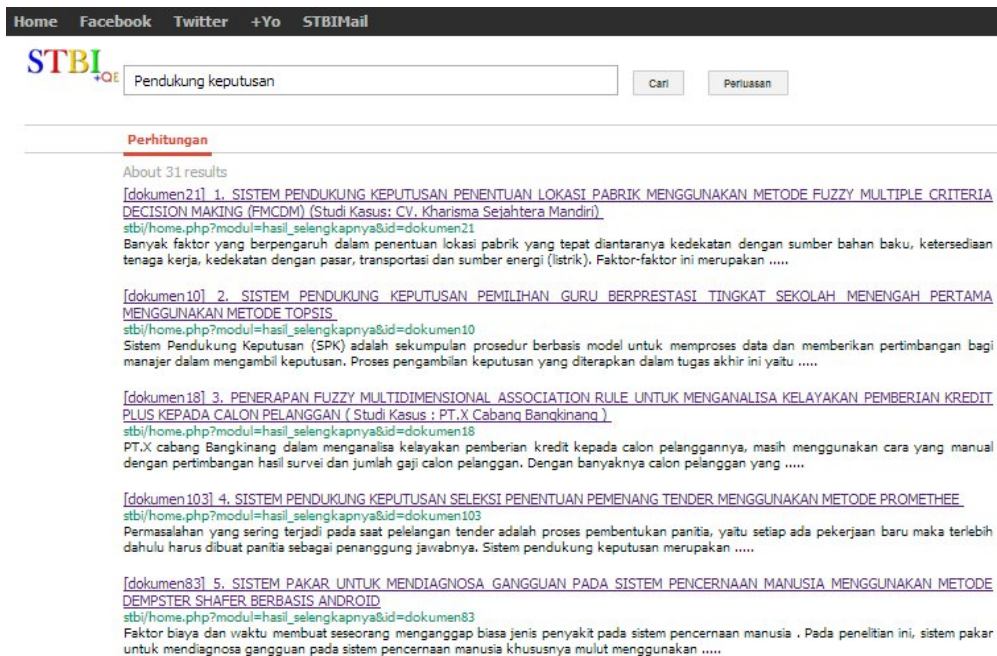
No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
1	103	Relevan	1/1= 1	1/15= 0.066667
2	10	Relevan	2/2= 1	2/15= 0.133333
3	16	Relevan	3/3= 1	3/15= 0.2
4	82	Relevan	4/4= 1	4/15= 0.266667
5	4	Relevan	5/5= 1	5/15= 0.333333
6	119	Relevan	6/6= 1	6/15= 0.4
7	87	Relevan	7/7= 1	7/15= 0.466667
8	97	Relevan	8/8= 1	8/15= 0.533333
9	113	Relevan	9/9= 1	9/15= 0.6
10	21	Relevan	10/10= 1	10/15= 0.666667
11	85	Relevan	11/11= 1	11/15= 0.733333

Tabel 5.7 Lanjutan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recal</i> (R)
12	88	Relevan	12/12= 1	12/15= 0.8
13	80	Relevan	13/13= 1	13/15= 0.866667
14	20	Relevan	14/14= 1	14/15= 0.933333
15	18	Relevan	15/15= 1	15/15= 1
16	3	Tidak Relevan	15/15= 0.9375	15/15= 1
17	115	Tidak Relevan	15/15= 0.8823529	15/15= 1
18	105	Tidak Relevan	15/15= 0.8333333	15/15= 1
19	62	Tidak Relevan	15/15= 0.7894737	15/15= 1
20	110	Tidak Relevan	15/15= 0.75	15/15= 1

### Pengujian *Query* 3 Dengan Perluasan

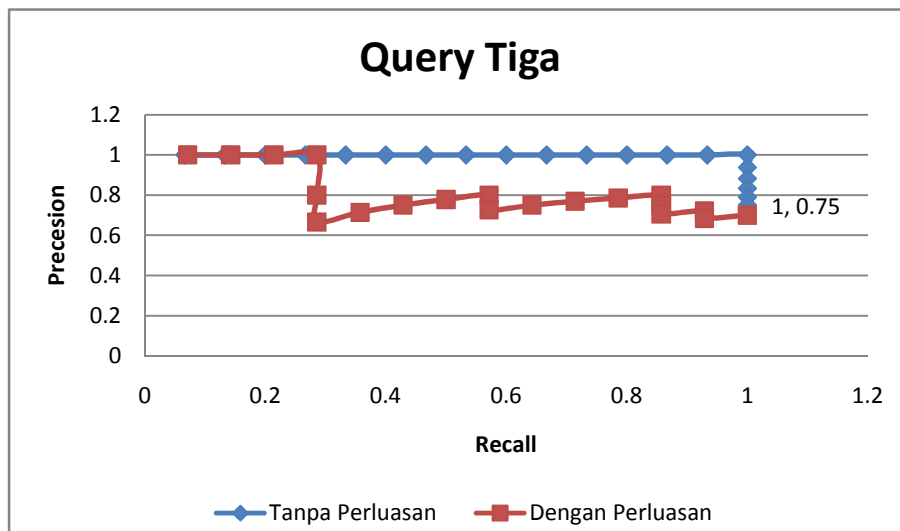
Pada pengujian dengan *input query* Q<sub>3</sub> (Pendukung keputusan) dengan Perluasa, maka *query* bertambah menjadi (Pendukung, penunjang) dan (Keputusan, kesimpulan pertimbangan ). Hasil pengujian dan perhitungan *recall* dan *presicion* selengkapnya dapat dilihat pada Gambar 5.22 dan Tabel 5.8

Gambar 5.22 Hasil Pengujian Q<sub>3</sub> Dengan Perluasan

Tabel 5.8 Perhitungan *recall* dan *Precision* ( $Q_3$ ) Dengan Perluasan

No	Id Dokumen	Relevan?	<i>Precision</i> (P)	<i>Recall</i> (R)
1	21	Relevan	1/1= 1	1/14= 0.07143
2	10	Relevan	2/2= 1	2/14= 0.14286
3	18	Relevan	3/3= 1	3/14= 0.21429
4	103	Relevan	4/4= 1	4/14= 0.28571
5	83	Tidak Relevan	4/5= 0.8	4/14= 0.28571
6	120	Tidak Relevan	4/6= 0.66667	4/14= 0.28571
7	20	Relevan	5/7= 0.71429	5/14= 0.35714
8	116	Relevan	6/8= 0.75	6/14= 0.42857
9	82	Relevan	7/9= 0.77778	7/14= 0.5
10	4	Relevan	8/10= 0.8	8/14= 0.57143
11	11	Tidak Relevan	8/11= 0.72727	8/14= 0.57143
12	119	Relevan	9/12= 0.75	9/14= 0.64286
13	87	Relevan	10/13= 0.76923	10/14= 0.71429
14	97	Relevan	11/14= 0.78571	11/14= 0.78571
15	113	Relevan	12/15= 0.8	12/14= 0.85714
16	17	Tidak Relevan	12/16= 0.75	12/14= 0.85714
17	34	Tidak Relevan	12/17= 0.70588	12/14= 0.85714
18	85	Relevan	13/18= 0.72222	13/14= 0.92857
19	114	Tidak Relevan	13/19= 0.68421	13/14= 0.92857
20	88	Relevan	14/20= 0.7	14/14= 1

Data interpolasi *recall* dan *precision* pada Tabel 5.7 dan 5.8 dapat digambarkan secara grafik garis sebagai berikut.



Gambar 5.23 Interpolasi *Recall* dan *Precision* Q<sub>3</sub> Tanpa perluasan

Dimana pada Gambar 5.23 diatas terlihat bahwa nilai *Precision* pada pengujian tanpa perluasan, lebih stabil jika dibanding dengan pengujian dengan perluasan. Contohnya pada titik kelima pada pencarian tanpa perluasan nilai *precision* samadengan 1 sedangkan hasil pencarian dengan perluasan bernilai 0,8. begitu juga dengan titik ke limabelas, dimana nilai *precision* tanpa perluasan samadengan 1 sedangkan dengan perluasan bernilai 0,8. Untuk perbandingan selengkapnya bisa langsung dilihat pada Tabel 5.7, Tabel 5.8 dan juga pada Gambar 5.23

Berikut pengujian yang tidak menggunakan *precision* yang tidak terinterpolasi dimana pada pengujian akan terlihat jumlah dokumen yang di-*retrieve* baik yang relevan maupun tidak dan jumlah dokumen yang tidak ter-*retrieve* baik yang relevan maupun tidak, dimana data tersebut akan digunakan untuk menghitung nilai *precision* dan *recall* berdasarkan masing-masing *query*.

Tabel 5.9. Hasil Pengujian *precision*(*P*) dan *recall*(*R*) pada Q<sub>1</sub> Tanpa Perluasan

	Relevan	Tidak Relevan
Ditemukan	19 ( <i>tp</i> )	42 ( <i>fp</i> )
Tidak ditemukan	0 ( <i>fn</i> )	59 ( <i>tn</i> )

Berdasarkan Tabel 5.9 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query* (*tp*) sebanyak 19 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 42 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query* (*fn*) sebanyak 0 dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 59 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 1 adalah:

$$P = tp / (tp + fp) = 19 / (19+42) = 19/61 = 0.31$$

$$R = tp / (tp + fn) = 19 / (19+0) = 19/19 = 1$$

Tabel 5.10. Hasil Pengujian *presicion*(*P*) dan *recall*(*R*) pada Q<sub>1</sub> Dengan Perluasan

	Relevan	Tidak Relevan
<b>Ditemukan</b>	19 ( <i>tp</i> )	65( <i>fp</i> )
<b>Tidak ditemukan</b>	0 ( <i>fn</i> )	36 ( <i>tn</i> )

Berdasarkan Tabel 5.10 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query* (*tp*) sebanyak 19 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 65 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query* (*fn*) sebanyak 0 dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 36 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 1 dengan perluasan adalah:

$$P = tp / (tp + fp) = 19/ (19+65) = 19/84 = 0.22$$

$$R = tp / (tp + fn) = 19 / (19+0) = 19/19 = 1$$

Tabel 5.11 Hasil Pengujian *presicion*(*P*) dan *recall*(*R*) pada Q<sub>2</sub> Tanpa Perluasan

	Relevan	Tidak Relevan
<b>Ditemukan</b>	32 ( <i>tp</i> )	9 ( <i>fp</i> )
<b>Tidak ditemukan</b>	0 ( <i>fn</i> )	79 ( <i>tn</i> )

Berdasarkan Tabel 5.11 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query* (*tp*) sebanyak 32 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 9 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query* (*fn*) sebanyak 0

dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 79 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 2 adalah:

$$P = tp / (tp + fp) = 32 / (32+9) = 32/41 = 0.78$$

$$R = tp / (tp + fn) = 32 / (32+0) = 32/32 = 1$$

Tabel 5.12 Hasil Pengujian *presicion(P)* dan *recall(R)* pada Q<sub>2</sub> Dengan Perluasan

	Relevan	Tidak Relevan
<b>Ditemukan</b>	32 ( <i>tp</i> )	10 ( <i>fp</i> )
<b>Tidak ditemukan</b>	0( <i>fn</i> )	74 ( <i>tn</i> )

Berdasarkan Tabel 5.12 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query* (*tp*) sebanyak 32 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 10 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query* (*fn*) sebanyak 0 dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 74 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 2 dengan perluasan adalah:

$$P = tp / (tp + fp) = 32 / (32+10) = 32/42 = 0.76$$

$$R = tp / (tp + fn) = 34 / (34+0) = 34/34 = 1$$

Tabel 5.13. Hasil Pengujian *presicion(P)* dan *recall(R)* pada Q<sub>3</sub> Tanpa Perluasan

	Relevan	Tidak Relevan
<b>Ditemukan</b>	16 ( <i>tp</i> )	9 ( <i>fp</i> )
<b>Tidak ditemukan</b>	0 ( <i>fn</i> )	95( <i>tn</i> )

Berdasarkan Tabel 5.13 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query* (*tp*) sebanyak 16 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 9 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query* (*fn*) sebanyak 0 dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 95 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 1 adalah:

$$P = tp / (tp + fp) = 16 / (16+9) = 16/25 = 0.64$$

$$R = tp / (tp + fn) = 16 / (16+0) = 16/16 = 1$$



Tabel 5.14. Hasil Pengujian *presicion(P)* dan *recall(R)* pada Q3 Dengan Perluasan

	Relevan	Tidak Relevan
Ditemukan	16 ( <i>tp</i> )	15 ( <i>fp</i> )
Tidak ditemukan	0 ( <i>fn</i> )	89 ( <i>tn</i> )

Berdasarkan Tabel 5.14 ditunjukkan bahwa jumlah dokumen yang dikembalikan yang relevan dengan *query (tp)* sebanyak 16 dokumen, sedangkan dokumen yang tidak relevan (*fp*) sebanyak 15 dokumen. Dan untuk jumlah dokumen yang tidak dikembalikan yang relevan dengan *query (fn)* sebanyak 0 dokumen, sedangkan dokumen yang tidak relevan sebanyak (*tn*) 89 dokumen. Maka, nilai *precision* dan *recall* untuk *query* 3 dengan perluasan adalah:

$$P = tp / (tp + fp) = 16 / (16+15) = 16/31 = 0.51$$

$$R = tp / (tp + fn) = 16 / (16+0) = 16/16 = 1$$

#### 5.2.4 Kesimpulan Pengujian Unjuk Kerja Sistem

Hasil pengujian yang diperoleh dari sistem temu kembali informasi yang menggunakan model ruang vektor dengan menerapkan Perluasan *query* sebagai berikut :

1. Persentase kualitas temu kembali informasi yang terjadi pada Q<sub>1</sub> terhadap jumlah dokumen yang di temu balikkan oleh sistem tanpa Perluasan *query* yaitu *precision* 31% dan *recall* 100%. Sedangkan dengan menerapkan perluasan menghasilkan *precision* 22% dan *recall* 100%.
2. Persentase kualitas temu kembali informasi yang terjadi pada Q<sub>2</sub> terhadap jumlah dokumen yang di temu balikkan oleh sistem tanpa Perluasan *query* yaitu *precision* 78% dan *recall* 100%. Sedangkan dengan menerapkan perluasan menghasilkan *precision* 76% dan *recall* 100%.
3. Persentase kualitas temu kembali informasi yang terjadi pada Q<sub>3</sub> terhadap jumlah dokumen yang di temu balikkan oleh sistem tanpa Perluasan *query* yaitu *precision* 64% dan *recall* 100%. Sedangkan dengan menerapkan perluasan menghasilkan *precision* 51% dan *recall* 100%.

## BAB VI

### PENUTUP

Pada bab ini akan diuraikan beberapa kesimpulan dari hasil yang didapatkan selama penelitian dan saran yang dapat digunakan pada penelitian selanjutnya.

#### 6.1 Kesimpulan

Setelah menyelesaikan tahapan-tahapan penelitian mengenai *query expansion* pada sistem temu kembali informasi dengan model ruang vektor, maka dapat diambil beberapa kesimpulan, yaitu :

1. Berdasarkan penelitian yang telah dilakukan, *Query expansion* pada sistem temu kembali informasi, ternyata belum cukup untuk meningkatkan performa dari hasil pencari yang ada, hal ini dikarenakan hasil dari *query expansion* lebih sering mengembalikan dokumen yang tidak relevan.
2. *Query expansion* cenderung menghasilkan dokumen yang tidak relevan dan bahkan bisa membuat dokumen yang tidak relevan terletak pada rengking pertama, meskipun telah diberikan faktor penyesuaian untuk menghindari hal tersebut.
3. *Query expansion* tidak dapat melakukan perluasan pada *query* yang menggunakan bahasa asing.
4. Pengujian sistem menggunakan *recall* dan *precision*, setiap pengguna memiliki sudut pandang yang berbeda dalam menilai kerelevanan dokumen terhadap *query*, hal ini dapat menyebabkan nilai yang berbeda pada pengguna yang berbeda.

## 6.2 Saran

Berdasarkan penelitian yang telah dilakukan, adapun saran-saran yang dapat dilakukan untuk perbaikan dan pengembangan sistem temu balik informasi mendatang, yaitu :

1. Untuk kata perluasan dari *query expansion* bisa ditambahkan dari kamus sinonim kata bahasa inggris atau yang dikenal dengan *WordNet*.
2. Koleksi dokumen yang digunakan sebagai korpus hendaklah memiliki kecendrungan bahasa yang tidak terlalu formal dan memiliki variasi bahasa yang banyak, contohnya seperti koleksi dokumen pada surat kabar.

## DAFTAR PUSTAKA

- Baeza, Yates R. dan Ribeiro, Neto B. "*Modern Information Retrieval*". Addison-Wesley. 1999.
- Budianita, Elvia, Studi temu kembali informasi (*information retrieval*) dengan model ruang vektor, pekanbaru : UIN suska riau, 2008
- Bunyamin, Hendra., dan chathalea, Puspa N." Aplikasi *Information Retrieval* (IR) CATA Dengan Metode *Generalized Vector Space Model*" *Jurnal Informatika*, Vol.4, No.1, hlm 29 – 38.2008
- Cios, krzysztof J, *Data Mining A Knowledge Discovery Approach*, new York : springer science, 2007
- Elvina, Irma., kudang, boro S., dan firman, ardiyansyah." desain konseptual penggunaan hyperlink sebagai alat bantu temu kembali informasi di perpustakaan," *Jurnal Perpustakaan Pertanian* Vol. 18, Nomor 1, hlm 14-23, 2009
- Halim, albert, perancangan dan pembuatan search engine dokumen menggunakan metode topic based vector space model, Surabaya: UK Petra, 2010
- Harjono, kristopher D., "perluasan vektor pada metode search vektor," *INTEGRAL*, vol.10, no.2 hlm.106-113, 2005
- Hariyono, Muhammad EA., dan wahyudi." customer information gathering menggunakan metode temu kembali informasi dengan model ruang vektor" *SNATI* vol-G, hlm G25-G28.2009
- Imran, Hazra., dan Aditi, Sharan "thesaurus and query expansion," *International Journal of Computer science & Information Technology (IJCSIT)*, Vol 1, No 2, hlm. 89-97, 2009"
- Mandala, Rila., dan Hendra, Stiwan : *Improving Information Retrieval System Performance by Automatic Query Expansion*. Bandung: ITB, 2002
- Manning, Christopher D, Ragnavan, Prabhakar, Schutze, Hinrich *Introduction to Information Retrieval*, Cambridge : Cambridge University Press, 2008.
- Metzler, Donald., Susan, dumais., dan Meek, Christopher." *Similarity Measures for Short Segments of Text*, "ECIR 2007 LNCS 4425, hlm 16-27, 2007
- Murad, azmi., Masrah, Azmi., dan Trevor, Martin. "Word Similarity for Document Grouping using Soft Computing," *IJCSNS International Journal of Computer Science and Network Security*, Vol.7, hlm.20-28, 2007

- Nugroho, Susetyo A.” Query expansion dengan menggabungkan metode Ruang vektor dan ordnet pada system Information retrieval” *Jurnal Informatika* Vol.5 No. 1, 2009
- Noor, agus M Y, analisa informasi retrieval system dengan model ruang vektor, bandung : UNIKOM,2011
- Polettini, Nicola., “*The Vector Space Model in Information Retrieval – Term Weighting Problem*,” *Department of Information and Communication Technology*, hlm 1-9,2004
- Ramadhany, Taufik. (2008). “*Implementasi Kombinasi Model Ruang Vektor dan Model Probabilistik Pada Sistem Temu Balik Informasi*.” Skripsi Terpublikasi. Bandung : Institut Teknologi Bandung.
- Salton, Gerard. *Introduction to Modern Information Retrieval*, New york : McGraw Hill, 1983.
- Sasongko, jati.” Aplikasi untuk Membangun Corpus dari Data Hasil *Crawling* dengan Berbagai Format Data Secara Otomatis,” *Jurnal Teknologi Informasi DINAMIK, Volume XV*,hlm.16-26,2010
- Sugono, dendy, thesaurus bahasa Indonesia pusat bahasa, Jakarta : DEPDIKNAS, 2008
- Syahroni,R, sistem temu balik informasi (STBI) laporan kerja praktek dan tugas akhir menggunakan model ruang vektor, pekanbaru: UIN suska Riau 2012
- Qiu Y. dan Frei H.P. “*Concept-based query expansion*”, in SIGIR ,1993.
- Zaenab, Ratu S.” efektivitas temu kembali informasi dengan menggunakan bahasa alami pada CD-ROM agribisnis dan cab abstrak” *Jurnal perpustakaan pertanian* Vol.11 No.2, hlm 41-49.2002

## DAFTAR RIWAYAT HIDUP

### Informasi Pribadi

#### Foto



#### Identitas

Nama : Khairi Lestari  
Tempat Lahir : Sei.Bela  
Tanggal Lahir : 10 : 02 : 1990  
Jenis Kelamin : Laki-Laki  
Agama : Islam  
Anak Ke : 2 Dari 5 Bersaudara  
Tinggi Badan : 173 cm

#### Keterangan

### Alamat

Sekarang : Jl-Garuda Sakti Gg Garuda 1  
No HP : 0852 6406 3331  
E-mail : Kherilest@yahoo.com

### Riwayat Pendidikan

1. Tahun 1996-2002 : MIS Nurul Iman Sungai Lokan
2. Tahun 2002-2005 : MTS Nurul Iman Sungai Lokan
3. Tahun 2005-2008 : SMA N 1 Tembilahan Hulu
4. Tahun 2008-2013 : Teknik Informatika UIN Suska Riau

## LAMPIRAN A

### PENGUJIAN KERELEVANAN

Pengujian Kerelevanan Hasil Pencarian *Query Expansion* Pada Sistem Temu Kembali Infomasi Dengan Model Ruang Vektor Berdasarkan Pendapat Para Ahli

#### Ahli Yang ditunjuk :

1. Benny Sukma Negara,ST,MT
2. Rizkqa Raaiqa B,ST
3. Sonya Meitracie,ST

#### *Query* yang diinputkan :

Tabel A.1 *Query* Pengujian

Nomor Pengujian	<i>Query</i> Tanpa Perluasan	<i>Query</i> dengan Perluasan
1	Keamanan Data	Keamanan Data Keselamatan Informasi
2	Komunikasi Mobile	Komunikasi mobile koneksi
3	Pendukung keputusan	Pendukung keputusan penunjang kesimpulan pertimbangan

#### Jawaban yang diberikan :

- R** = Untuk Hasil yang relevan  
**TR** = Untuk Hasil yang tidak relevan  
**KET** = Untuk kesimpulan dari pendapat ketiga ahli

## 1. Pengujian 1 Tanpa Perluasan

Tabel A.2 Hasil Pencarian menggunakan *query* 1 Tanpa Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	66	PERANCANGAN SISTEM KEAMANAN DALAM PENTRANSMISIAN DATA DARI TAG MENUJU READER PADA RFID	R	R	R	<b>R</b>
2	64	IMPLEMENTASI METODE ENKRIPSI RC5 UNTUK KEAMANAN DATA PADA KOMUNIKASI PERANGKAT BERGERAK	R	R	R	<b>R</b>
3	56	PERANCANGAN DAN PEMBUATAN SISTEM ENKRIPSI -DEKRIPSI DATA IMAGE DENGAN METODE CHAOS	R	R	R	<b>R</b>
4	19	PENERAPAN ALGORITMA RC4 DAN HASHING MD5 UNTUK MENSIMULASIKAN KUNCI ELEKTRONIK PADA HANDPHONE MELALUI BLUETOOTH	R	R	R	<b>R</b>
5	71	APLIKASI KOMPRESI BERBASIS HUFFMAN DAN ENKRIPSI BERBASIS AES PADA PENGIRIMAN SMS	R	R	R	<b>R</b>
6	55	APLIKASI ENKRIPSI SMS MENGGUNAKAN METODE BLOWFISH	R	R	R	<b>R</b>
7	65	IMPLEMENTASI VIDEO PENGAMAT DAN PENGAMANAN VIDEO MENGGUNAKAN ALGORITMA VIDEO ENKRIPSI (VEA)	R	R	R	<b>R</b>
8	70	PENGEMBANGAN SISTEM PERPARKIRAN MENGGUNAKAN ENKRIPSI DATA DAN TEKNOLOGI BARCODE	R	R	R	<b>R</b>
9	98	APLIKASI KRIPTOGRAFI SEBAGAI PENGAMANAN DATA PADA FILE DENGAN ALGORITMA BLOWFISH	R	R	R	<b>R</b>
10	68	KAJIAN SISTEM PENGAMAN DATA YANG MENJAMIN KEASLIAN DAN KERAHASIAAN DATA DENGAN ALGORITMA RSA	R	R	R	<b>R</b>
11	72	ENKRIPSI SMS MENGGUNAKAN ECC (ELLIPTIC CURVE CRYPTOGRAPHY)	R	R	R	<b>R</b>
12	57	ENKRIPSI CITRA JPEG2000 TANPA MENGHASILKAN MARKER CODE	R	R	R	<b>R</b>
13	73	PEMANFAATAN METODE VIDEO ENKRIPSI PADA VIDEO STREAMING UNTUK PERLINDUNGAN HAK CIPTA	R	R	R	<b>R</b>
14	52	IMPLEMENTASI ALGORITMA RIPEMD-160 SEBAGAI FUNGSI PENYANDI PROSES OTENTIKASI PASSWORD PADA ENKRIPSI DATA BERBASIS DES	R	R	R	<b>R</b>



15	92	ANALISIS PERFORMANSI ALGORITMA RC6 DENGAN PENDEKATAN PANJANG KUNCI DAN JUMLAH ROUND	R	R	R	<b>R</b>
16	109	RANCANG BANGUN APLIKASI PENJUALAN BUKU BERBASIS WEB YANG MENERAPKAN REKOMENDASI HASIL PENCARIAN MENGGUNAKAN DATA MINING KLASIFIKASI	TR	R	R	<b>R</b>
17	69	KOMPRESI DAN ENKRIPSI SMS DENGAN METODE HUFFMAN CODE DAN ALGORITMA ENIGMA	R	R	R	<b>R</b>
18	61	ANALISIS ALGORITMA ENKRIPSI ELGAMAL, GRAIN V1, DAN AES DENGAN STUDI KASUS APLIKASI RESEP MASAKAN DI ANDROID	R	R	R	<b>R</b>
19	63	IMPLEMENTASI ENKRIPSI/DEKRIPSI SINYAL WICARA PADA DSP TMS320C5402 BERBASIS RSA DAN RC4	R	R	R	<b>R</b>
20	107	PERANCANGAN SISTEM INFORMASI PENGOLAHAN DATA PELANGGAN PADA PERUSAHAAN DAERAH AIR MINUM (PDAM) KOTA PEKANBARU	TR	TR	R	<b>TR</b>

## 2. Pengujian 1 dengan Perluasan

Tabel A.3 Hasil Pencarian menggunakan *query* 1 Dengan Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	107	PERANCANGAN SISTEM INFORMASI PENGOLAHAN DATA PELANGGAN PADA PERUSAHAAN DAERAH AIR MINUM (PDAM) KOTA PEKANBARU	TR	TR	R	<b>TR</b>
2	30	LOCATION BASED ALERT PADA MOBILE PHONE BERBASIS ANDROID	TR	R	R	<b>R</b>
3	34	PEMBUATAN APLIKASI MOBILE CLIENT TRAFFIC REPORT SYSTEM BERBASIS SISTEM OPERASI ANDROID	TR	TR	R	<b>TR</b>
4	19	PENERAPAN ALGORITMA RC4 DAN HASHING MD5 UNTUK MENSIMULASIKAN KUNCI ELEKTRONIK PADA HANDPHONE MELALUI BLUETOOTH	R	R	R	<b>R</b>
5	111	SISTEM TEMU BALIK INFORMASI (STBI) LAPORAN KERJA PRAKTEK DAN TUGAS AKHIR MENGGUNAKAN MODEL RUANG VEKTOR (STUDI KASUS : TEKNIK INFORMATIKA)	TR	TR	R	<b>TR</b>
6	114	RANCANG BANGUN APLIKASI MOBILE KULINER (eMKUL) MENGGUNAKAN LOCATION BASED SERVICE (LBS) BERBASIS ANDROID	TR	TR	R	<b>TR</b>
7	17	STUDI TEMU KEMBALI INFORMASI (INFORMATION RETRIEVAL) DENGAN MODEL RUANG VEKTOR	TR	TR	TR	<b>TR</b>
8	78	TUGAS AKHIR LAYANAN SISTEM INFORMASI BURSA DAN PEREKRUTAN TENAGA KERJA PERUSAHAAN MENGGUNAKAN TEKNOLOGI WAP ( Wireless Application Protocol)	TR	TR	TR	<b>TR</b>

9	108	RANCANG BANGUN SISTEM INFORMASI PEMASANGAN DAN PENAYANGAN IKLAN DI RTV DENGAN KONSEP CUSTOMER RELATIONSHIP MANAGEMENT	TR	TR	TR	<b>TR</b>
10	2	PERANCANGAN DAN IMPLEMENTASI SISTEM INFORMASI GEOGRAFIS KOTA PEKANBARU BERBASIS MOBILE MENGGUNAKAN J2ME DAN GOOGLE STATIC MAPS API	TR	TR	TR	<b>TR</b>
11	66	PERANCANGAN SISTEM KEAMANAN DALAM PENRANSMISIAN DATA DARI TAG MENUJU READER PADA RFID	R	R	R	<b>R</b>
12	84	RANCANG BANGUN APLIKASI PHP ENCODER DENGAN MENGGUNAKAN ALGORITMA RC4	TR	TR	TR	<b>TR</b>
13	12	RANCANG BANGUN SISTEM INFORMASI UNTUK PEMILIHAN MEREK PAKAIAN DENGAN ANALISIS MARKOV (Studi Kasus : PT.DELAMI)	TR	TR	TR	<b>TR</b>
14	31	APLIKASI BULLETIN BOARD MOBILE BERBASIS ANDROID	TR	TR	TR	<b>TR</b>
15	56	PERANCANGAN DAN PEMBUATAN SISTEM ENKRIPSI -DEKRIPSI DATA IMAGE DENGAN METODE CHAOS	R	R	R	<b>R</b>
16	52	APLIKASI EMAIL CLIENT PADA MOBILE PHONE MENGGUNAKAN TEKNOLOGI BLUETOOTH	TR	TR	TR	<b>TR</b>
17	64	IMPLEMENTASI METODE ENKRIPSI RC5 UNTUK KEAMANAN DATA PADA KOMUNIKASI PERANGKAT BERGERAK	R	R	R	<b>R</b>
18	90	PERANCANGAN DAN IMPLEMENTASI SISTEM INFORMASI PEMBAYARAN SPP DI SMA N 1 BANGKINANG BARAT	TR	TR	TR	<b>TR</b>
19	70	PENGEMBANGAN SISTEM PERPARKIRAN MENGGUNAKAN ENKRIPSI DATA DAN TEKNOLOGI BARCODE	R	R	TR	<b>R</b>
20	65	IMPLEMENTASI VIDEO PENGAMAT DAN PENGAMANAN VIDEO MENGGUNAKAN ALGORITMA VIDEO ENKRIPSI (VEA)	R	R	R	<b>R</b>

### 3. Pengujian 2 Tanpa Perluasan

Tabel A.4 Hasil Pencarian menggunakan *query* 2 Tanpa Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	36	PERANCANGAN DAN PENGEMBANGAN APLIKASI MOBILE LEARNING BERBASIS JAVA UNTUK PENYEDIAAN DAN PENYUNTINGAN BAHAN AJAR	R	R	R	<b>R</b>
2	54	GENERATOR ANGKA ACAK DALAM GAME SUDOKU DENGAN METODE BACKTRACKING PADA ANDROID OS MOBILE	R	R	R	<b>R</b>

3	39	PENERAPAN PEMROGRAMAN JAVA PADA MOBILE DEVICE ENVIRONMENT	R	R	R	<b>R</b>
4	43	STUDI SIMULASI PERFORMANSI MOBILE IPV4 (MIPV4) DAN MOBILE IPV6 (MIPV6)	R	R	R	<b>R</b>
5	100	RANCANG BANGUN APLIKASI MOBILE LEARNING CLIENT SERVER BERBASIS MOODLE PADA PLATFORM ANDROID	R	R	R	<b>R</b>
6	52	APLIKASI EMAIL CLIENT PADA MOBILE PHONE MENGGUNAKAN TEKNOLOGI BLUETOOTH	R	R	R	<b>R</b>
7	41	KINERJA SISTEM KOMUNIKASI KOOPERATIF DENGAN MODEL KANAL MOBILE-TO-MOBILE DOUBLE RING STATISTIK	R	R	R	<b>R</b>
8	49	KINERJA SISTEM KOMUNIKASI KOOPERATIF DENGAN KANAL MOBILE-TO-MOBILE Rician FADING	R	R	R	<b>R</b>
9	33	PEMBANGUNAN SISTEM PELACAKAN DAN PENELUSURAN DEVICE MOBILE BERBASIS GLOBAL POSITIONING SISTEM (GPS) PADA PLATFORM MOBILE GOOGLE ANDROID	R	R	R	<b>R</b>
10	44	RANCANG BANGUN PERANGKAT LUNAK VOICE MAIL UNTUK MOBILE DEVICE BERBASIS SYMBIAN OS MENGGUNAKAN TEKNOLOGI BLUETOOTH	R	R	R	<b>R</b>
11	48	RANCANG BANGUN PERANGKAT LUNAK MESIN PENCARI FILE PDF UNTUK PERANGKAT MOBILE	R	R	R	<b>R</b>
12	37	RANCANG BANGUN PERANGKAT LUNAK MESIN PENCARI FILE PDF UNTUK PERANGKAT MOBILE	R	R	R	<b>R</b>
13	38	APLIKASI GEOMETRI MOBILE LEARNING UNTUK TINGKAT TK MENGGUNAKAN TEKNOLOGI JAVA 2 MICRO EDITION	R	R	R	<b>R</b>
14	31	APLIKASI BULLETIN BOARD MOBILE BERBASIS ANDROID	R	R	R	<b>R</b>
15	45	KINERJA KOMUNIKASI KOOPERATIF MENGGUNAKAN SISTEM MIMO DENGAN MODEL KANAL MOBILE-TO-MOBILE TWO-ROSE-RING	R	R	R	<b>R</b>
16	32	IMPLEMENTASI APLIKASI MOBILE CHAT PADA SISTEM OPERASI ANDROID 2.2	R	R	R	<b>R</b>
17	47	KAJIAN EMISI CO2 MENGGUNAKAN PERSAMAAN MOBILE 6 DAN MOBILE COMBUSTION DARI SEKTOR TRANSPORTASI DI KOTA SURABAYA	R	R	TR	<b>R</b>
18	30	LOCATION BASED ALERT PADA MOBILE PHONE BERBASIS ANDROID	R	R	R	<b>R</b>
19	64	IMPLEMENTASI METODE ENKRIPSI RC5 UNTUK KEAMANAN DATA PADA KOMUNIKASI PERANGKAT BERGERAK	R	TR	TR	<b>TR</b>
20	40	MOBILE AGENT UNTUK MENENTUKAN WAKTU SHOLAT DENGAN MENGGUNAKAN AGLET	R	R	R	<b>R</b>

#### 4. Pengujian 2 dengan perluasan

Tabel A.5 Hasil Pencarian menggunakan *query* 2 Dengan Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	53	RANCANG BANGUN APLIKASI BERPINDAH PENGENDALI ROBOT BERBASIS ANDROID MENGGUNAKAN KONEKSI BLUETOOTH	R	R	R	<b>R</b>
2	32	IMPLEMENTASI APLIKASI MOBILE CHAT PADA SISTEM OPERASI ANDROID 2.2	R	R	TR	<b>R</b>
3	19	PENERAPAN ALGORITMA RC4 DAN HASHING MD5 UNTUK MENSIMULASIKAN KUNCI ELEKTRONIK PADA HANDPHONE MELALUI BLUETOOTH	R	TR	R	<b>R</b>
4	36	PERANCANGAN DAN PENGEMBANGAN APLIKASI MOBILE LEARNING BERBASIS JAVA UNTUK PENYEDIAAN DAN PENYUNTINGAN BAHAN AJAR	R	R	TR	<b>R</b>
5	114	RANCANG BANGUN APLIKASI MOBILE KULINER (eMKUL) MENGGUNAKAN LOCATION BASED SERVICE (LBS) BERBASIS ANDROID	R	R	TR	<b>R</b>
6	54	GENERATOR ANGKA ACAK DALAM GAME SUDOKU DENGAN METODE BACKTRACKING PADA ANDROID OS MOBILE	R	R	TR	<b>R</b>
7	39	PENERAPAN PEMROGRAMAN JAVA PADA MOBILE DEVICE ENVIRONMENT	R	R	TR	<b>R</b>
8	43	STUDI SIMULASI PERFORMANSI MOBILE IPV4 (MIPV4) DAN MOBILE IPV6 (MIPV6)	R	R	TR	<b>R</b>
9	100	RANCANG BANGUN APLIKASI MOBILE LEARNING CLIENT SERVER BERBASIS MOODLE PADA PLATFORM ANDROID	R	R	R	<b>R</b>
10	52	APLIKASI EMAIL CLIENT PADA MOBILE PHONE MENGGUNAKAN TEKNOLOGI BLUETOOTH	R	R	R	<b>R</b>
11	41	KINERJA SISTEM KOMUNIKASI KOOPERATIF DENGAN MODEL KANAL MOBILE-TO-MOBILE DOUBLE RING STATISTIK	TR	R	R	<b>R</b>
12	49	KINERJA SISTEM KOMUNIKASI KOOPERATIF DENGAN KANAL MOBILE-TO-MOBILE RICIAN FADING	TR	R	R	<b>R</b>
13	33	PEMBANGUNAN SISTEM PELACAKAN DAN PENELUSURAN DEVICE MOBILE BERBASIS GLOBAL POSITIONING SISTEM (GPS) PADA PLATFORM MOBILE GOOGLE ANDROID	R	R	TR	<b>R</b>
14	44	RANCANG BANGUN PERANGKAT LUNAK VOICE MAIL UNTUK MOBILE DEVICE BERBASIS SYMBIAN OS MENGGUNAKAN TEKNOLOGI BLUETOOTH	R	R	R	<b>R</b>
15	48	RANCANG BANGUN PERANGKAT LUNAK MESIN PENCARI FILE PDF UNTUK PERANGKAT MOBILE	R	R	TR	<b>R</b>

16	37	APLIKASI GEOMETRI MOBILE LEARNING UNTUK TINGKAT TK MENGGUNAKAN TEKNOLOGI JAVA 2 MICRO EDITION	R	R	TR	<b>R</b>
17	38	PERANCANGAN DAN PEMBUATAN PERANGKAT LUNAK VIDEO VIEWER BERBASIS BLUETOOTH PADA MOBILE PHONE	R	R	R	<b>R</b>
18	31	APLIKASI BULLETIN BOARD MOBILE BERBASIS ANDROID	R	R	TR	<b>R</b>
19	45	KINERJA KOMUNIKASI KOOPERATIF MENGGUNAKAN SISTEM MIMO DENGAN MODEL KANAL MOBILE-TO-MOBILE TWO-ROSE-RING	TR	R	R	<b>R</b>
20	47	KAJIAN EMISI CO2 MENGGUNAKAN PERSAMAAN MOBILE 6 DAN MOBILE COMBUSTION DARI SEKTOR TRANSPORTASI DI KOTA SURABAYA	TR	R	TR	<b>TR</b>

### 5. Pengujian 3 Tanpa Perluasan

Tabel A.5 Hasil Pencarian menggunakan *query* 3 Tanpa Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	103	SISTEM PENDUKUNG KEPUTUSAN SELEKSI PENENTUAN PEMENANG TENDER MENGGUNAKAN METODE PROMETHEE	R	R	R	<b>R</b>
2	10	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN GURU BERPRESTASI TINGKAT SEKOLAH MENENGAH PERTAMA MENGGUNAKAN METODE TOPSIS	R	R	R	<b>R</b>
3	16	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN BARU MENGGUNAKAN METODE AHP DAN TOPSIS	R	R	R	<b>R</b>
4	82	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN LOKASI PENGASPALAN JALAN MENGGUNAKAN FUZZY MADM DENGAN METODE SAW	R	R	R	<b>R</b>
5	4	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN BARU MENGGUNAKAN METODE AHP DAN TOPSIS (STUDI KASUS : PT. FARISHINDO PERKASA RAYA)	R	R	R	<b>R</b>
6	119	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN MAHASISWA KEPERAWATAN PRAKTEK TERBAIK DENGAN METODE DECISION TREE DAN METODE PERBANDINGAN EKSPONENSIAL (MPE)	R	R	R	<b>R</b>
7	87	SISTEM PENDUKUNG KEPUTUSAN MENENTUKAN PENERIMA BEA SISWA MENGGUNAKAN FMADM (FUZZY MULTIPLE ATTRIBUTE DECISION MAKING) DAN WP (WEIGHTED PRODUCT)	R	R	R	<b>R</b>
8	97	SISTEM PENDUKUNG KEPUTUSAN UNTUK PEMILIHAN DAN PENEMPATAN KARYAWAN BARU DENGAN MODEL PREFERENCE RANKING ORGANIZATION METHOD FOR	R	R	R	<b>R</b>

		ENRICHMENT EVALUATION (PROMETHEE)				
9	113	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN PEMBANGUNAN LABORATORIUM KOMPUTER SEKOLAH DI KABUPATEN KEPULAUAN MERANTI MENGGUNAKAN (SAW)	R	R	R	<b>R</b>
10	21	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN LOKASI PABRIK MENGGUNAKAN METODE FUZZY MULTIPLE CRITERIA DECISION MAKING (FMCDM)	R	R	R	<b>R</b>
11	85	SISTEM PENDUKUNG KEPUTUSAN PRIORITAS NASABAH UNTUK MENDAPATKAN KREDIT PEMILIKAN RUMAH (KPR) MENGGUNAKAN METODE FUZZY AHP (F-AHP) DAN TOPSIS	R	R	R	<b>R</b>
12	88	SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PENEMPATAN SALES PADA SUATU AREA PEMUKIMAN DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR (KNN)	R	R	R	<b>R</b>
13	80	PENERAPAN METODE K-NEAREST NEIGHBOR DALAM RANCANG BANGUN SISTEM PENDUKUNG KEPUTUSAN REKOMENDASI DIET BAGI PENDERITA PENYAKIT DIABETES MELLITUS GESTASIONAL	R	R	R	<b>R</b>
14	20	PENDEKATAN ANALYTICAL HIERARCHY PROCESS (AHP) UNTUK ALAT BANTU KEPUTUSAN PEMILIHAN REKANAN PROYEK	R	R	R	<b>R</b>
15	18	PENERAPAN FUZZY MULTIDIMENSIONAL ASSOCIATION RULE UNTUK MENGANALISA KELAYAKAN PEMBERIAN KREDIT PLUS KEPADA CALON PELANGGAN	R	R	R	<b>R</b>
16	3	IMPLEMENTASI ALGORITMA APRIORI UNTUK MENGANALISIS KERANJANG BELANJA KONSUMEN PADA DATA TRANSAKSI PENJUALAN SUPERMARKET	TR	TR	TR	<b>TR</b>
17	115	RANCANG BANGUN APLIKASI PENGELOMPOKAN SASARAN BISNIS DENGAN METODE CLUSTERING C-MEANS	TR	TR	TR	<b>TR</b>
18	105	PENGGUNAKAN METODE BRANCH AND BOUND UNTUK MENYELESAIKAN MASALAH CAPITAL BUDGETING	TR	TR	TR	<b>TR</b>
19	62	IMPLEMENTASI ALGORITMA RIPEMD-160 SEBAGAI FUNGSI PENYANDI PROSES OTENTIKASI PASSWORD PADA ENKRIPSI DATA BERBASIS DES	TR	TR	TR	<b>TR</b>
20	110	APLIKASI MARKET BASKET ANALYSIS DENGAN MENGGUNAKAN ALGORITMA APRIORI DAN ALGORITMA FREQUENT PATTERN GROWTH-SQL BASED FREQUENT PATTERN MINING	TR	TR	TR	<b>TR</b>

## 6. Pengujian 3 Dengan Perluasan

Tabel A.7 Hasil Pencarian menggunakan *query* 3 Dengan Perluasan

No	Id Dok	Judul Dokumen	Ahli 1	Ahli 2	Ahli 3	Ket :
1	21	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN LOKASI PABRIK MENGGUNAKAN METODE FUZZY MULTIPLE CRITERIA DECISION MAKING (FMCDM)	R	R	R	<b>R</b>
2	10	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN GURU BERPRESTASI TINGKAT SEKOLAH MENENGAH PERTAMA MENGGUNAKAN METODE TOPSIS	R	R	R	<b>R</b>
3	18	PENERAPAN FUZZY MULTIDIMENSIONAL ASSOCIATION RULE UNTUK MENGANALISA KELAYAKAN PEMBERIAN KREDIT PLUS KEPADA CALON PELANGGAN	R	R	R	<b>R</b>
4	103	SISTEM PENDUKUNG KEPUTUSAN SELEKSI PENENTUAN PEMENANG TENDER MENGGUNAKAN METODE PROMETHEE	R	R	R	<b>R</b>
5	83	SISTEM PAKAR UNTUK MENDIAGNOSA GANGGUAN PADA SISTEM PENCERNAAN MANUSIA MENGGUNAKAN METODE DEMPSTER SHAFER BERBASIS ANDROID	TR	TR	TR	<b>TR</b>
6	120	SISTEM PAKAR UNTUK MENENTUKAN GANGGUAN PERKEMBANGAN PADA ANAK DENGAN METODE CERTAINTY FACTOR	TR	TR	TR	<b>TR</b>
7	20	PENDEKATAN ANALYTICAL HIERARCHY PROCESS (AHP) UNTUK ALAT BANTU KEPUTUSAN PEMILIHAN REKANAN PROYEK	R	R	R	<b>R</b>
8	116	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN BARU MENGGUNAKAN METODE AHP DAN TOPSIS	R	R	R	<b>R</b>
9	82	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN LOKASI PENGASPALAN JALAN MENGGUNAKAN FUZZY MADM DENGAN METODE SAW	R	R	R	<b>R</b>
10	4	SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN BARU MENGGUNAKAN METODE AHP DAN TOPSIS (STUDI KASUS : PT. FARISHINDO PERKASA RAYA)	R	R	R	<b>R</b>
11	11	ENKRIPSI SMS (SHORT MESSAGE SERVICE) MENGGUNAKAN ALGORITMA RC6 PADA SISTEM OPERASI ANDROID	TR	TR	TR	<b>TR</b>
12	119	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN MAHASISWA KEPERAWATAN PRAKTEK TERBAIK DENGAN METODE DECISION TREE DAN METODE PERBANDINGAN EKSPONENSIAL (MPE)	R	R	R	<b>R</b>
13	87	SISTEM PENDUKUNG KEPUTUSAN MENENTUKAN PENERIMA BEA SISWA MENGGUNAKAN FMADM (FUZZY MULTIPLE ATTRIBUTE DECISION MAKING) DAN WP	R	R	R	<b>R</b>
14	97	SISTEM PENDUKUNG KEPUTUSAN UNTUK PEMILIHAN DAN PENEMPATAN KARYAWAN	R	R	R	<b>R</b>

		BARU DENGAN MODEL PREFERENCE RANKING ORGANIZATION METHOD FOR ENRICHMENT EVALUATION (PROMETHEE)				
15	113	SISTEM PENDUKUNG KEPUTUSAN PENENTUAN PEMBANGUNAN LABORATORIUM KOMPUTER SEKOLAH DI KABUPATEN KEPULAUAN MERANTI MENGGUNAKAN METODE SIMPLE ADDITIVE WEIGHTING (SAW)	R	R	R	<b>R</b>
16	17	STUDI TEMU KEMBALI INFORMASI (INFORMATION RETRIEVAL) DENGAN MODEL RUANG VEKTOR	TR	TR	TR	<b>TR</b>
17	34	PEMBUATAN APLIKASI MOBILE CLIENT TRAFFIC REPORT SYSTEM BERBASIS SISTEM OPERASI ANDROID	TR	TR	TR	<b>TR</b>
18	85	SISTEM PENDUKUNG KEPUTUSAN PRIORITAS NASABAH UNTUK MENDAPATKAN KREDIT PEMILIKAN RUMAH (KPR) MENGGUNAKAN METODE FUZZY AHP (F-AHP) DAN TOPSIS	R	R	R	<b>R</b>
19	114	RANCANG BANGUN APLIKASI MOBILE KULINER (eMKUL) MENGGUNAKAN LOCATION BASED SERVICE (LBS) BERBASIS ANDROID	TR	TR	TR	<b>TR</b>
20	88	SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PENEMPATAN SALES PADA SUATU AREA PEMUKIMAN DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR (KNN)	R	R	R	<b>R</b>

Kesimpulan :

Berdasarkan pendapat dari tiga ahli diatas dapat ditarik kesimpulan sebagai berikut :

1. Dari hasil pengujian Satu pencarian tanpa perluasan menghasilkan 19 dokumen relevan dan dengan perluasan menghasilkan 8 dokumen relevan.
2. Dari hasil pengujian Satu pencarian tanpa perluasan menghasilkan 19 dokumen relevan dan dengan perluasan menghasilkan 19 dokumen relevan.
3. Dari hasil pengujian Satu pencarian tanpa perluasan menghasilkan 15 dokumen relevan dan dengan perluasan menghasilkan 14 dokumen relevan.
4. Rata-rata dari tiga pengujian pencarian tanpa perluasan menemukan 88% dokumen relevan dan dengan perluasan 68% dokumen relevan